

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2001年6月28日 (28.06.2001)

PCT

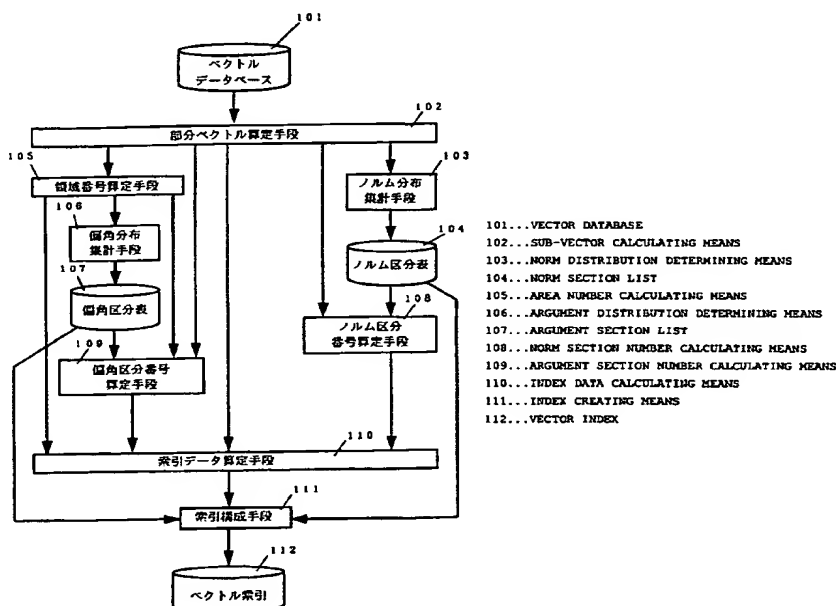
(10) 国際公開番号
WO 01/46858 A1

- (51) 国際特許分類: G06F 17/30 (KANNO, Yuji) [JP/JP]; 〒232-0061 神奈川県横浜市南区大岡1丁目1-19-504 Kanagawa (JP).
- (21) 国際出願番号: PCT/JP00/09079
- (22) 国際出願日: 2000年12月21日 (21.12.2000) (74) 代理人: 二瓶正敬(NIHEI, Masayuki); 〒160-0004 東京都新宿区四谷2丁目12-5 第6富沢ビル6F Tokyo (JP).
- (25) 国際出願の言語: 日本語 (81) 指定国 (国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (26) 国際公開の言語: 日本語
- (30) 優先権データ: 特願平 11/363058 1999年12月21日 (21.12.1999) JP (84) 指定国 (広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (71) 出願人 (米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒571-0050 大阪府門真市大字門真1006番地 Osaka (JP).
- (72) 発明者: および
- (75) 発明者/出願人 (米国についてのみ): 菅野祐司

[続葉有]

(54) Title: VECTOR INDEX CREATING METHOD, SIMILAR VECTOR SEARCHING METHOD, AND DEVICES FOR THEM

(54) 発明の名称: ベクトル索引作成方法と類似ベクトル検索方法およびそれらのための装置



(57) Abstract: Similar vectors are fast retrieved from a vector database of several hundreds of dimensions with reference to a single vector index according to the criterion of the inner product or distance, after specifying the similarity search range and the maximum number of similar vectors to be retrieved. For creating the vector index,

[続葉有]



添付公開書類:
— 国際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

each vector is decomposed into sub-vectors and featured by a norm section, an assigned area, and an argument section. For similarity search, a sub-query vector and a sub search range are determined from the query vector and the search range, similarity search in sub-space is carried out, and differences from the search range are cumulated to determine the upper limits. An accurate criterion having a higher upper limit is preferentially determined, thereby producing a final similarity search result.

(57) 要約:

本発明は、数百次元のベクトルデータベースから、単一のベクトル索引によって、内積あるいは距離のいずれかの尺度による、類似検索範囲と最大取得件数を指定した類似ベクトル検索を高速に行うもので、ベクトル索引作成の際には、各ベクトルを複数の部分ベクトルに分解し、ノルム区分と所属領域と偏角区分とによって特徴付けることで索引を作成し、類似検索の際には、質問のベクトルと検索範囲とから部分質問ベクトルと部分検索範囲を求めて、各部分空間で類似検索を行って検索範囲との差分を累算して上限値を求め、上限値の高いものから正確な尺度を求めていって最終的な類似検索結果を得る。

明 細 書

ベクトル索引作成方法と類似ベクトル検索方法およびそれらのための装置

5 技術分野

- 本発明は、テキスト情報、画像情報、音声情報、アンケート結果、売り上げ（POS）データなど、電子化されて蓄積された各種のデータベース（データの集合）中の各データの特徴抽出を行って作成したベクトルデータ（当該データの特徴を表す実数をN個並べた、通常「特徴ベクトル」と呼ばれるN次元実ベクトルと、その識別番号などとの組）の集合であるベクトルデータベースに対して、計算機を利用してベクトルデータの検索、分類、傾向分析などを行うための索引の作成方法およびその装置と、指定したベクトルに類似のベクトルを、上記の方法および装置により作成された索引を用いて効率よく検索する、類似ベクトル検索方法およびその装置に関する。

背景技術

- 近年、テキスト、画像、音声などのマルチメディア情報のデータベース化や、POSシステムなどの普及に伴い、マルチメディアデータベースシステムやデータマイニングシステムなどのコンピュータシステムにおいて、数十～数百次元のベクトルデータを数十万個～数百万個集めたベクトルデータベースの検索、分類、傾向分析などを効率よく実行するための技術の研究開発が盛んになっている。

- 例えば新聞記事データベースの場合には、多数の新聞記事データを蓄積したデータベースに対し、W語の単語を集めた辞書を使って各新聞記事から辞書中の各単語kの出現回数 f_k を抽出し、各新聞記事を、識別番

号 i と、 W 次元の実ベクトル (f_1, f_2, \dots, f_w) の組で表現する。主成分分析手法などによってこのベクトルを変換し、各ベクトルの主要な N 個 ($N \ll W$) の成分を求めて、これをベクトルデータとする。

- 指定した新聞記事に対応するベクトルデータと、データベース中の他の
- 5 新聞記事に対応するベクトルとの内積を計算し、内積の最も大きいベクトルを持つ新聞記事を求めることで、高精度な類似記事検索が可能になる。米国特許第 4 8 3 9 8 5 3 号には、このような、ベクトルデータを用いた文書検索方法が開示されている。

- また、写真データベースの場合には、多数の写真画像のデータを蓄積
- 10 したデータベースに対し、各写真データに対して 2 次元のフーリエ変換を施し、主要な N 個のフーリエ成分を f_k を抽出し、各写真データを、写真番号 i と、 N 次元の実ベクトル (f_1, f_2, \dots, f_w) の組で表現し、これをベクトルデータとする。指定した写真に対応するベクトルデータと、データベース中の他の写真データに対応するベクトルとの
- 15 距離 (2 つのベクトルの差の大きさ) を計算し、距離の最も小さいベクトルを持つ写真データを求めることで、高精度な類似写真検索が可能になる。さらに、例えば「ポートレート」, 「風景写真」, 「花の接写」といった、異なるカテゴリーに属する典型的な写真データをそれぞれ数枚ずつ、分類条件として提示し、各カテゴリーの平均的な特徴ベクトル
- 20 を算出、各写真データのベクトルに対して最も距離の小さい特徴ベクトルのカテゴリーを割り当てることで、残りの写真データを上記の 3 種類のカテゴリーに自動的に分類することが可能になる。

- このような用途には、数十～数百次元という、極めて高次元のベクトルの効率のよい類似検索手法が必要となるため、各種の方法が研究され
- 25 ている。例えば、Norio Katayama and Shinichi Satoh, “The SR-tree: An Index Structure for High-Dimensional Nearest Neighbor Queries”,

Proceedings of the SIGMOD ' 97, ACM (1997) には、SR (多次元探索) 木を用いた高次元ベクトルの索引作成方法および類似検索方法が開示されている。また、Sergey Brin, "Near Neighbor Search in Large Metric Spaces", Proceedings of the VLDB ' 95, Morgan-Kaufman Publishers
5 (1995) には、ボロノイ分割に基づいた高次元ベクトルの索引作成方法および類似検索方法が開示されている。さらに、Stefan Berchtold, Christian Bohm and Hans Kriegel, "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality", Proceedings of the SIGMOD ' 98, ACM (1998) には、「ピラミッド技法」というデータ分割
10 手法に基づいた高次元ベクトルの索引作成方法および類似検索方法が開示されている。

しかしながら、これらの従来のベクトル索引作成方法および類似ベクトル検索方法は、以下の4種の条件のいずれかが満たされず、広範囲の用途に広く適用できない、という課題があった。

- 15 1) ベクトルの次元が数百次元の場合でも、高速な検索ができる。
- 2) 類似検索の際に、ベクトル間の距離と、ベクトルの内積の2種類の類似度のいずれかを選択できる。
- 3) 「もっとも類似しているL個のベクトルを求める」という型の類似検索ができる。さらに、Lが比較的大きい(数十~数百)場合でも、
20 検索処理が極端に遅くならない。
- 4) 「内積が0.6以上」といった、類似検索範囲の指定ができる。
- 5) 索引作成に要する計算量が実用的範囲である(即ち、ベクトルデータ量nに比例した時間、あるいは $n * \log(n)$ 時間で索引が作成できる)。

25 具体的には、SR木を用いた方法では上記1)、2)が満たされず、ボロノイ分割に基づいた方法では2)、5)が満たされず、ピラミッド

技法を用いた方法では、2)、3)が満たされない。

本発明のベクトル索引作成方法および類似ベクトル検索方法およびその装置は、これらの従来の技術の持つ課題を解決するもので、高次元のベクトルを、複数個の部分ベクトルに分解し、各部分ベクトルの方向と
5 大きさを、中心ベクトルで規定した所属領域番号と、中心ベクトルとのなす角（偏角）と、ノルムを表すノルム区分との組で表現して記録することで、どのような質問ベクトルに対しても、ベクトル索引の検索対象範囲を精度よく限定することができ、部分内積の下限値（部分二乗距離の上限値）と、実際の部分内積（部分二乗距離）との差を累算すること
10 で、分岐限定技法による効率的な検索結果の確定ができるため、上記の1)～4)をすべて満足した、広範囲の用途に適用可能なベクトル索引作成方法および類似ベクトル検索方法を提供するものである。

上記課題を解決するために、本発明の第1の様態によるベクトル索引作成方法および装置は、部分ベクトルを算定する手段と、ノルム分布を
15 集計してノルム区分表を作成する手段と、領域番号を算定する手段と、偏角分布を集計して偏角区分表を作成する手段と、ノルム区分番号を算定する手段と、偏角区分番号を算定する手段と、索引データを算定する手段と、索引を構成する手段とを有することにより、方向、ノルムの分布が不明なベクトルデータベースに対して、ベクトルの次元が数百次元
20 の場合でも、高速な検索が可能で、類似検索の際に、ベクトル間の距離と、ベクトルの内積の2種類の類似度のいずれかを選択でき、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「内積が0.6以上」といった、類似検索範囲の指定が可能
25 で、索引作成に要する計算量が実用的範囲であるような、ベクトル索引の作成を可能にする、という効果を奏するものである。

また、本発明の第2の様態によるベクトル索引作成方法および装置は、第1の様態に加えて、成分区分番号を算定する手段をさらに有することにより、第1の様態による効果に加えて、成分の量子化による計算誤差を最小限に抑えながら、作成すべきベクトル索引の容量を大幅に低減できる、という効果を奏するものである。

また、本発明の第3の様態による類似ベクトル検索方法および装置は、部分質問条件を算定する手段と、検索対象範囲を生成する手段と、索引を検索する手段と、内積差分上限を算定する手段と、類似検索結果を決定する手段とを有し、部分内積差分の累算値を算定して類似検索の手がかりとして用いることにより、ベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「内積が0.6以上」といった、類似検索範囲の指定が可能で、内積を類似尺度として用いる類似ベクトル検索を可能にする、という効果を奏するものである。

また、本発明の第4の様態による類似ベクトル検索方法および装置は、部分質問条件を算定する手段と、検索対象範囲を生成する手段と、索引を検索する手段と、二乗距離差分上限を算定する手段と、類似検索結果を決定する手段とを有し、部分二乗距離差分の累算値を算定して類似検索の手がかりとして用いることにより、ベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「距離が0.8以下」といった、類似検索範囲の指定が可能で、距離を類似尺度として用いる類似ベクトル検索を可能にする、と

いう効果を奏するものである。

図面の簡単な説明

- 図 1 は、第 1 の実施の形態におけるベクトル索引作成装置の全体構成
5 を示すブロック図であり、
- 図 2 は、第 2 の実施の形態におけるベクトル索引作成装置の全体構成
を示すブロック図であり、
- 図 3 は、第 3 の実施の形態における類似ベクトル検索装置の全体構成
を示すブロック図であり、
- 10 図 4 は、第 4 の実施の形態における類似ベクトル検索装置の全体構成
を示すブロック図であり、
- 図 5 A、5 B は一体となって、第 1、第 2 の実施の形態におけるベク
トル索引作成の第 1 ステップの作成手順を記述した流れ図であり、
- 図 6 A、6 B は一体となって、第 1 の実施の形態におけるベクトル索
15 引作成の第 2、第 3 ステップの作成手順を記述した流れ図であり、
- 図 7 A、7 B は一体となって、第 2 の実施の形態におけるベクトル索
引作成の第 2、第 3 ステップの作成手順を記述した流れ図であり、
- 図 8 A、8 B は一体となって、第 3 の実施の形態における類似ベクト
ル検索の第 1 ステップの検索手順を記述した流れ図であり、
- 20 図 9 は、第 3 の実施の形態における類似ベクトル検索の第 2 ステップ
の検索手順を記述した流れ図であり、
- 図 10 A、10 B は一体となって、第 4 の実施の形態における類似ベ
クトル検索の第 1 ステップの検索手順記述した流れ図であり、
- 図 11 A、11 B は一体となって、第 4 の実施の形態における類似ベ
25 クトル検索の第 2 ステップの検索手順を記述した流れ図であり、
- 図 12 A、12 B は一体となって、第 1、第 2、第 3、第 4 の実施の

形態におけるベクトルデータベースの内容例を示す一覧図であり、

図 1 3 は、第 1、第 2 の実施の形態におけるノルム分布集計結果例を示す特性図であり、

図 1 4 は、第 1、第 2 の実施の形態における偏角分布集計結果例を示す特性図であり、

図 1 5 A、1 5 B は一体となって、第 1、第 2、第 3、第 4 の実施の形態におけるノルム区分表の内容例を示す一覧図であり、

図 1 6 は、第 1、第 2、第 3、第 4 の実施の形態における偏角区分表の内容例を示す一覧図であり、

10 図 1 7 A、1 7 B は一体となって、第 3 の実施の形態における表 W の内容例（一部分）を示す一覧図であり、

図 1 8 A、1 8 B、1 8 C は一体となって、第 4 の実施の形態における表 W の内容例（一部分）を示す一覧図である。

15 発明を実施するための最良の形態

<第 1 の実施の形態>

以下、本発明の第 1 の実施の形態について、図面を参照しながら説明する。

20 （ベクトル索引作成装置の構成）

図 1 は本発明の請求項 1、3～8、14、16～21 にかかる、ベクトル索引作成装置の第 1 の実施の形態の全体構成を表わすブロック図である。図 1 において、ベクトルデータベース 101 は、20 万件の新聞記事を集めた新聞記事全文データベースから作成した、各新聞記事の特徴を表す 296 次元の単位実ベクトルと、1 から 200000 の範囲の識別番号の 2 項目からなるベクトルデータを 20 万個格納し、図 1 2 A、

1 2 Bのような内容を持つ。

部分ベクトル算定手段 1 0 2 は、ベクトルデータベース 1 0 1 中の各ベクトルデータの 2 9 6 次元のベクトル V に対し、3 7 種の 8 次元部分ベクトル $v_0 \sim v_{36}$ と、0 ~ 3 6 の部分空間番号 b とを算定する。

- 5 ノルム分布集計手段 1 0 3 は、部分ベクトル算定手段 1 0 2 が 2 0 万個の各ベクトルデータについて算定した 3 7 個の各部分ベクトルのユークリッドノルムを算出して、その分布を集計し、2 5 6 個の連続する実数範囲、

ノルム区分 0 = [0 , r_1) ,

- 10 ノルム区分 1 = [r_1 , r_2) ,

...

ノルム区分 2 5 5 = [r_{255} , r_{256})

としてノルム区分を定める。

- 15 ノルム区分表 1 0 4 は、ノルム分布集計手段 1 0 3 が算出したノルム区分を格納する。

領域番号算定手段 1 0 5 は、部分ベクトル算定手段 1 0 2 が算定した 8 次元の各部分ベクトル v に対して、成分が { 0 , 1 , - 1 } のいずれかであるような、0 ベクトルでない 8 次元ベクトルをノルムが 1 になるよう正規化した、

- 20 領域中心ベクトル 0 = (0 , 0 , 0 , 0 , 0 , 0 , 0 , 1) ,

領域中心ベクトル 1 = (0 , 0 , 0 , 0 , 0 , 0 , 0 , - 1) ,

領域中心ベクトル 2 = (0 , 0 , 0 , 0 , 0 , 0 , 1 , 0) ,

領域中心ベクトル 3 = $\text{sqrt}(1/2) * (0, 0, 0, 0, 0, 0, 1, 1)$,

- 25 領域中心ベクトル 4 = $\text{sqrt}(1/2) * (0, 0, 0, 0, 0, 0, 1, -1)$,

領域中心ベクトル 5 = (0, 0, 0, 0, 0, 0, -1, 0),

...

領域中心ベクトル 6 5 5 4 = $\text{sqrt}(1/7) * (-1, -1, -1, -1, -1, -1, 1, 0)$,

5 領域中心ベクトル 6 5 5 5 = $\text{sqrt}(1/8) * (-1, -1, -1, -1, -1, -1, 1, 1)$,

領域中心ベクトル 6 5 5 6 = $\text{sqrt}(1/8) * (-1, -1, -1, -1, -1, -1, 1, -1)$,

10 領域中心ベクトル 6 5 5 7 = $\text{sqrt}(1/7) * (-1, -1, -1, -1, -1, -1, 0, 0)$,

領域中心ベクトル 6 5 5 8 = $\text{sqrt}(1/8) * (-1, -1, -1, -1, -1, -1, 1, 1)$,

領域中心ベクトル 6 5 5 9 = $\text{sqrt}(1/8) * (-1, -1, -1, -1, -1, -1, -1, -1)$,

15 という (ただし「 $\text{sqrt}(x)$ 」は x の平方根を表す)、6 5 6 0 個のベクトルを領域中心ベクトルとし、部分ベクトル v との内積が最も大きい領域中心ベクトル p_d を求めて、その番号 d を v の所属領域の領域番号とし、 p_j と v とのなす角の余弦を、偏角 c として求める。

20 偏角分布集計手段 1 0 6 は、領域番号算定手段 1 0 5 が 2 0 万個の各ベクトルデータの 3 7 個の部分ベクトルについて算定した偏角の値 c の分布を集計し、4 個の連続する実数範囲、

偏角区分 0 = [c_0, c_1),

偏角区分 1 = [c_1, c_2),

偏角区分 2 = [c_2, c_3),

25 偏角区分 3 = [c_3, c_4),

として偏角区分を定める。

偏角区分表 107 は、偏角分布集計手段 106 が算出した偏角区分を格納する。

ノルム区分番号算定手段 108 は、部分ベクトル算定手段 102 が算出した各部分ベクトルのノルムの属するノルム区分番号 r を、ノルム区分表 104 を検索して定める。

偏角区分番号算定手段 109 は、部分ベクトル算定手段 102 が算出した各部分ベクトル v と、領域番号算定手段 105 が v に対して算定した領域中心ベクトル p から、 v と p の偏角の属する偏角区分番号 c を、偏角区分表 107 を検索して定める。

10 索引データ算定手段 110 は、部分ベクトル算定手段 102 の算定した部分ベクトル v_b および部分空間番号 b と、領域番号算定手段 105 の算定した領域番号 d と、偏角区分番号算定手段 109 の算定した偏角区分番号 c と、ノルム区分番号算定手段 108 の算定したノルム区分番号 r とから検索用の鍵、

15 $K = ((b * 6560 + d) * 4 + c) * 256 + r$

を作成し、この鍵 K と部分ベクトルの識別番号 i と成分 v_b の組 (K, i, v_b) を索引データとして算定する。

索引構成手段 111 は、索引データ算定手段 110 の算定した索引データ (K, i, v_b) から、 K を鍵とし、 (i, v_b) を検索するような探索木と、各識別番号 i と各部分空間番号 b の組に対して、領域番号 d と偏角区分番号 c と、ノルム区分番号 r とから第 2 の鍵、

$L = (d * 4 + c) * 256 + r$

を格納した逆探索表と、ノルム区分表 104 と、偏角区分表 107 とを格納した索引を構成する。

25 ベクトル索引 112 は、索引構成手段 111 が作成する探索木と逆探索表とノルム区分表 104 と偏角区分表 107 とを格納する。

(ベクトル索引作成装置の動作)

以上の構成のベクトル索引作成装置において、その動作を図面をもとに説明する。図5A、5Bは一体となってベクトル索引作成の第1ステップにおけるノルム区分表Rおよび偏角区分表Cの作成処理の手順を記述した流れ図であり、図6A、6Bは一体となってベクトル索引作成の第2、第3ステップにおける索引登録データの算定およびベクトル索引の作成処理の手順を記述した流れ図である。図中の「 $\text{sqrt}(x)$ 」は x の平方根を、「 $\text{int}(x)$ 」は x の整数部分を、「 $\text{abs}(x)$ 」は x の絶対値を、それぞれ表すものとする。また「 $\text{sign2}(x)$ 」は x が非負なら1を、負なら2をとる関数であるとする。

(ベクトル索引作成の第1ステップ)

ベクトル索引作成の第1ステップでは、まず、部分ベクトル算定手段102が、ベクトルデータベース101からベクトルデータを順に読み込んで部分ベクトルを算定し、ノルム分布集計手段103と偏角分布集計手段106が、それぞれ部分ベクトルのノルム分布、偏角分布を算定する。全てのベクトルデータを処理した時点で、ノルム区分表および偏角区分表が作成される。ベクトルデータベース中のベクトルのノルムの上限値は既知であり、上限値は r_sup であるとする。本実施の形態の例では、各ベクトルデータのベクトルが単位ベクトルなので、明らかに $r_sup = 1$ となる。ベクトルデータベース中のベクトルのノルムの上限値が未知である場合には、予め調査して r_sup を求めておけばよい。

初めに、ステップ1001で、集計用の表 H_r および H_c を0に初期化し、総部分ベクトル数 n も0にする。次に、ステップ1002で、未処理のベクトルデータ(i, v)をベクトルデータベースから1つ読み

込む。部分空間の番号 b を 0 に初期化する。ステップ 1 0 0 3 で、 b の値に従って、8 次元の部分ベクトル u を、読み込んだ 2 9 6 次元のベクトル v の先頭から、連続する 8 成分ずつに区分して 3 7 種作成する。例えば図 1 2 A の 1 番目のベクトルデータの場合には、 $b = 0$ の部分ベ

5 トルは、

(+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032
-0.007255 -0.007715)

となる。 $b = 1$ の部分ベクトルは、

(-0.025648 +0.016061 -0.060584 -0.013593 -0.020985 -0.112403
10 -0.012045 +0.044741)

となり、 $b = 3 6$ の部分ベクトルは、

(+0.069379 +0.020206 +0.032996 +0.047815 +0.046106 +0.001794
+0.035342-0.003895)

となる。そして、 u のノルム $|u|$ をノルムの最大値 r_sup で割って 1 0 0 0 0 倍した後、整数化し、ノルム分布集計表 H_r の該当する区
15 分 j に累算し、ノルム分布を集計する。

図 1 3 はこのようにして集計したノルム分布のグラフの例である。グラフの横軸はノルム分布集計表 H_r の区分番号、縦軸は各区分番号 j に対する $H_r[j]$ の値、即ち区分 j のノルム範囲のノルムを持つ部分ベ
20 クトルの数である。図 1 2 A の 1 番目のベクトルデータの $b = 0$ の部分ベクトルの場合には、

$$|u| = \sqrt{(0.029259^2 + 0.016005^2 + \dots + 0.007715^2)} = 0.049193$$

であり、 $r_sup = 1$ であるから、区分 j は、

$$25 \quad j = \text{int}((0.049193 / 1.0) * 10000) = 491$$

となる。

偏角区分は、ステップ1004～ステップ1009で集計する。まず、ステップ1004で、部分ベクトル u の8個の成分 $u[0] \sim u[7]$ に対して、絶対値の大きい順に、成分の番号を格納する。図12Aの1番目のベクトルデータの $b=0$ の部分ベクトルの場合には、第0成分の絶対値が最も大きく、次に第3成分の絶対値が大きく、第4成分の絶対値が最も小さいので、

$$s[0..7] = (0 \ 3 \ 2 \ 1 \ 5 \ 7 \ 6 \ 4)$$

となる。

次にステップ1005～ステップ1008を、変数 m の値を0から7まで変えながら、8回（8＝部分空間の次元）繰り返し、6560種の領域中心ベクトルのうちで最も部分ベクトル u との内積が大きいものの番号 d と、そのときの内積の値 x とを求める。ステップ1005では、絶対値の大きいほうから $m+1$ 個の成分が、（部分ベクトルの成分の符号） $\times 1$ で、残りの $7-m$ 個の成分が0であるような領域中心ベクトルの番号 j と内積の \sqrt{m} 倍の値 y とを求めている。ステップ1006ではステップ1005で求めた値 y から内積を $y \times \sqrt{1/m}$ で計算し、これまでの内積の最高値 x と比較し、 x より大きければステップ1007で内積の最高値 x と、領域中心ベクトルの番号 d とを更新する。このように、成分が $\{+1, 0, -1\}$ のいずれかであるような領域中心ベクトル群を用いることにより、非常に簡単な計算によって、部分ベクトルと最も内積の大きい領域中心ベクトルの番号と、その内積の値とを効率よく求めることができる。

図12Aの1番目のベクトルデータの $b=0$ の部分ベクトルの場合には、

$$(|u[0]|) \times \sqrt{1/1} = 0.029259$$

$$(|u[0]| + |u[3]|) \times \sqrt{1/2} = 0.038361$$

$$(|u[0]| + |u[3]| + |u[2]|) * \text{sqrt}(1/3) = 0.043514$$

$$(|u[0]| + |u[3]| + |u[2]| + |u[1]|) * \text{sqrt}(1/4) = 0.045687$$

$$(|u[0]| + |u[3]| + |u[2]| + |u[1]| + |u[5]|) * \text{sqrt}(1/5) = 0.044903$$

$$(|u[0]| + |u[3]| + |u[2]| + |u[1]| + |u[5]| + |u[7]|) * \text{sqrt}(1/6) =$$

5 0.044140

$$(|u[0]| + |u[3]| + |u[2]| + |u[1]| + |u[5]| + |u[7]| + |u[6]|) * \text{sqrt}(1/7) = 0.043608$$

$$(|u[0]| + |u[3]| + |u[2]| + |u[1]| + |u[5]| + |u[7]| + |u[6]| + |u[4]|) * \text{sqrt}(1/8) = 0.043217$$

10 となって、内積の最高値である $x = 0.045687$ 、および領域中心ベクトル

$$(+1/2, -1/2, -1/2, +1/2, 0, 0, 0, 0)$$

の番号である $d = (3^7) + 2 * (3^6) + 2 * (3^5) + (3^4) = 4212$ が求まる。

15 続いてステップ1009で内積 x を部分ベクトル u のノルムで割って、部分ベクトルと領域中心ベクトルとのなす角の余弦を求め、10000倍した後、整数化し、偏角分布集計表 H_c の該当する区分 j に累算し、偏角分布を集計する。図14はこのようにして集計した偏角分布のグラフの例である。グラフの横軸は偏角分布集計表 H_c の区分番号、縦軸は

20 各区分番号 j に対する $H_c[j]$ の値、即ち区分 j の偏角範囲の偏角を持つ部分ベクトルの数である。ただし図14において、8274より小さな区分の H_c の集計値は全て0なので、8000～10000区分の部分のみ図示している。図12Aの1番目のベクトルデータの $b = 0$ の部分ベクトルの場合には、

$$\begin{aligned} 25 \quad j &= \text{int}(10000 * 0.045687 / 0.049193) \\ &= \text{int}(10000 * 0.928730) = 9287 \end{aligned}$$

となる。

部分ベクトル選択用変数 b と、総部分ベクトル数集計用変数 n を増やした後、ステップ 1 0 1 0 で着目ベクトルデータの全ての部分ベクトルを処理したか否かを判定し、まだ未処理の部分ベクトルが残っている場合
5 にはステップ 1 0 0 3 に戻って、次の部分ベクトルに対して処理を行う。全ての部分ベクトルを処理した場合には、ステップ 1 0 1 1 でベクトルデータベース 1 0 1 中の全てのベクトルデータを処理したか否かを判定し、まだ未処理のベクトルデータが残っている場合にはステップ 1 0 0 2 に戻って次のベクトルデータを処理し、全てのベクトルデータを
10 読み込んで処理した場合には、ステップ 1 0 1 2 ～ステップ 1 0 1 8 のノルム区分表、偏角区分表の作成処理に進む。

ステップ 1 0 1 2 では作業変数を初期化し、ステップ 1 0 1 3 ～ステップ 1 0 1 8 でノルム区分表および偏角区分表の区分データの作成処理を行う。ステップ 1 0 1 3 では、ノルム集計結果のうち、ノルムが 0 から
15 $r_sup * j / 10000$ までの部分ベクトルの個数の合計値 x と、偏角集計結果のうち、偏角が 0 から $j / 10000$ までの部分ベクトルの個数の合計値 y とを求める。

ステップ 1 0 1 4 では、ノルムが 0 から $r_sup * j / 10000$ までの部分ベクトルの個数の全部分ベクトル数に対する割合 x / n が、
20 ノルム区分表の 2 5 6 個の区分のうち、第 k 区分までの区分数の割合 $k / 256$ より大きいか否かを判定する。大きい場合には、ステップ 1 0 1 5 に進んでノルム区分表の第 k 区分の境界値である $R[k]$ を $r_sup * j / 10000$ と定める。図 1 5 A、1 5 B は一体となって、以上のようにして図 1 3 のノルム分布のノルム分布集計表 H_r から作成した
25 ノルム区分表の例である。分布が集中している 0. 1 ～ 0. 2 の区間を細かく区分していることがわかる。

ステップ1016からステップ1017で、偏角区分についても、同様にして偏角区分表の第m区分の境界値を定める。ステップ1018でノルム集計結果、偏角集計結果をすべて処理したか否かを判定し、未処理の集計結果が残っている場合にはステップ1013に戻って処理を続
5 行し、全ての集計結果を処理し終えた場合にはステップ1019に進んで、 $R[0..256]$ 、 $C[0..4]$ をそれぞれノルム区分表、偏角区分表としてベクトル索引作成の第1ステップを終了する。図16は、以上のようにして図14の偏角分布の偏角分布集計表 H_c から作成した偏角区分表の例である。分布が集中している0.95付近を細かく区分していること
10 がわかる。

(ベクトル索引作成の第2ステップ)

ベクトル索引作成の第2ステップでは、図6A、6Bのステップ1101～ステップ1109に記載した処理を行って、個々の部分ベクトル
15 から索引登録データを作成する。まず、1101で探索木 T を初期化し、 T の登録データ数を0とする。探索木としては、

1) 整数値を鍵として、ベクトルデータ (i, u) 即ち整数と8個の浮動小数点数の組が登録できる。

2) 登録時の整数値の範囲を鍵にして、登録したデータを検索できる。
20 という、2つの条件を満たすものであれば、R. セジウィック著、野下浩平他訳、“アルゴリズム第2巻 探索・文字列・計算幾何”，近代科学社(1992)や、G. H. Gonnet著 玄光男他訳“アルゴリズムとデータ構造ハンドブック”，啓学出版(1987)などの教科書に記載されている、B木や2分探索木などの(平衡)探索木を用いることができる。

25 ステップ1102では、ベクトルデータベース101から未処理のベクトルデータを1つ読み込み、部分空間番号 b を0から順に増やしなが

ら各部分空間の部分ベクトルに対する処理を行う。ステップ1103では、部分ベクトル u を作成し、作成済みのノルム区分表104を検索して、ノルム $|u|$ に対応するノルム区分の番号 r を求める。ステップ1104～ステップ1108までは、図5A、5Bのステップ1004～

5 ステップ1008と同一の処理を行って、6560種の領域中心ベクトルのうちで最も部分ベクトル u との内積が大きいものの番号 d と、そのときの内積の値 x とを求める。

ステップ1109では、作成済みの偏角区分表107を検索して、偏角(即ち部分ベクトルと所属領域の領域中心ベクトルとのなす角の余弦)

10 $x/|u|$ に対応する偏角区分の番号 c を求める。ステップ1110では、索引データ算定手段110が、こうして求めたノルム区分番号 d 、偏角区分番号 c から、

$$\begin{aligned} k &= b * N_d * N_c * N_r + d * N_c * N_r + c * N_r + r \\ &= b * 7617440 + d * 1024 + c * 256 + r \end{aligned}$$

15 によって、部分空間番号 b 、領域番号 d 、偏角区分番号 c 、ノルム区分番号 r の4種の整数値を1つの整数値に変換し、探索木に登録する際の鍵 k を算定、ステップ1111でこの鍵 k と、部分ベクトルデータ(i , u)から、索引登録データ(k , i , u)を算定する。ただし、 N_d は領域の総数で6560、 N_c は偏角区分数で4、 N_r はノルム区分数で256である。このようにして、ベクトル索引作成の第2ステップで、各ベクトルデータの各部分ベクトルに対する索引登録データ(k , i , u)を効率よく(ベクトルデータ数に比例した時間で)作成できる。

20

(ベクトル索引作成の第3ステップ)

25 ベクトル索引作成の第3ステップでは、図6Bのステップ1111～ステップ1115に記載した処理を行って、索引登録データからベクトル

- ル索引を作成する。まず、ステップ 1 1 1 1 では、索引登録データ (k, i, u) 中の k を鍵として、データ (i, u) を探索木に (追加) 登録する。次にステップ 1 1 1 2 で、逆探索表 K の識別番号 i のベクトルデータの部分空間番号 b に対応する要素 K [i, u] に鍵 k を格納する。部分空間番号 b を 1 増やした後、ステップ 1 1 1 3 で全ての部分空間について処理を終了したか否かを判定する。まだ未処理の部分空間が残っている場合にはステップ 1 1 0 3 に戻って次の部分ベクトルに対する処理を行い、全ての部分空間について処理を終了した場合には、ステップ 1 1 1 4 に進む。ステップ 1 1 1 4 では、ベクトルデータベース 1 0 1 中の全てのベクトルデータを処理したか否かを判定する。まだ未処理のベクトルデータが残っている場合にはステップ 1 1 0 2 に戻って次のベクトルデータに対する処理を行い、全てのベクトルデータについて処理を終了した場合には、ステップ 1 1 1 5 に進んで探索木 T、逆探索表 K、ノルム区分表 R、偏角区分表 C を格納したベクトル索引を作成して、ベクトル索引の作成が完了する。

以上説明したように、本発明の第 1 の実施の形態のベクトル索引作成方法および装置によれば、以下のような優れた効果を奏する。

- 1) 2 9 6 次元のベクトルを、3 7 種の 8 次元の部分ベクトルに分解し、各々の部分ベクトルに対して、その方向を、6 5 6 0 種の領域中の所属領域の領域番号と偏角区分番号の組で精密に量子化し、その大きさをノルム区分番号で量子化した上で、複数の鍵を 1 つの整数値にエンコードして探索木に登録することで、部分空間毎の高速で高精度な範囲検索が可能となる。

- 2) また、逆探索表を作成・装備することで、ベクトルデータの識別番号を指定してベクトルの成分を取得する機能が、成分データを 2 重持ちすることなく実現できるため、元のベクトルデータベース 1 0 1 が検

索時には不要となり、検索装置の記憶容量を削減することが可能となる。

3) ノルム分布集計手段、偏角分布集計手段において、各区分に所属する部分ベクトルの個数ができる限り均一になるように区分境界を定めるため、分布に偏りのあるベクトルデータベースに対しても、常に最適な（検索速度の低下を最小限に抑えた）ベクトル索引を作成することができる。

4) 領域中心ベクトルとして、成分が $\{0, +1, -1\}$ のいずれかである 0 ベクトルでない全てのベクトルを正規化したベクトルの集合を用いているため、各部分ベクトルの所属領域の算定が、領域の個数には依存せずに行え、部分ベクトルの成分の絶対値順序の算出と、成分の絶対値の加算などの、極めて少ない計算量で済むことから、数十万～数百万個のベクトルデータからなる大規模なベクトルデータベースの場合でも、実用的な処理時間でベクトル索引を作成することが可能となる。

15 <第2の実施の形態>

次に、本発明の第2の形態について、図面を参照しながら説明する。

(ベクトル索引作成装置の構成)

図2は本発明の請求項2、3～8、15、16～21にかかる、ベクトル索引作成装置の第2の実施の形態の全体構成を表わすブロック図である。図2において、ベクトルデータベース201は、20万件の新聞記事を集めた新聞記事全文データベースから作成した、各新聞記事の特徴を表す296次元の単位実ベクトルと、1から200000の範囲の識別番号と、記事の見出しの3項目からなるベクトルデータを20万個格納し、図12A、12Bのような内容を持つ。

部分ベクトル算定手段202は、ベクトルデータベース201中の各

ベクトルデータの296次元のベクトルVに対し、37種の8次元部分ベクトル $v_0 \sim v_{36}$ と、0～36の部分空間番号bとを算定する。

ノルム分布集計手段203は、部分ベクトル算定手段202が20万個の各ベクトルデータについて算定した37個の各部分ベクトルのユー

- 5 クリッドノルムを算出して、その分布を集計し、256個の連続する実数範囲、

ノルム区分0 = $[0, r_1)$,

ノルム区分1 = $[r_1, r_2)$,

...

- 10 ノルム区分255 = $[r_{255}, r_{256})$,

としてノルム区分を定める。

ノルム区分表204は、ノルム分布集計手段203が算出したノルム区分を格納する。

- 領域番号算定手段205は、部分ベクトル算定手段202が算定した
15 8次元の各部分ベクトルvに対して、成分が{0, 1, -1}のいずれかであるような、0ベクトルでない8次元ベクトルをノルムが1になるよう正規化した、

領域中心ベクトル0 = (0, 0, 0, 0, 0, 0, 0, 1),

領域中心ベクトル1 = (0, 0, 0, 0, 0, 0, 0, -1),

- 20 領域中心ベクトル2 = (0, 0, 0, 0, 0, 0, 1, 0),

領域中心ベクトル3 = $\text{sqrt}(1/2) * (0, 0, 0, 0, 0, 0, 1, 1)$,

領域中心ベクトル4 = $\text{sqrt}(1/2) * (0, 0, 0, 0, 0, 0, 1, -1)$,

- 25 領域中心ベクトル5 = (0, 0, 0, 0, 0, 0, -1, 0),

...

領域中心ベクトル 6 5 5 4 = $\text{sqr t} (1/7) * (-1, -1, -1, -1, -1, -1, 1, 0)$,

領域中心ベクトル 6 5 5 5 = $\text{sqr t} (1/8) * (-1, -1, -1, -1, -1, -1, 1, 1)$,

5 領域中心ベクトル 6 5 5 6 = $\text{sqr t} (1/8) * (-1, -1, -1, -1, -1, -1, 1, -1)$,

領域中心ベクトル 6 5 5 7 = $\text{sqr t} (1/7) * (-1, -1, -1, -1, -1, -1, 0)$,

10 領域中心ベクトル 6 5 5 8 = $\text{sqr t} (1/8) * (-1, -1, -1, -1, -1, -1, 1)$,

領域中心ベクトル 6 5 5 9 = $\text{sqr t} (1/8) * (-1, -1, -1, -1, -1, -1, -1)$,

という (ただし「 $\text{sqr t} (x)$ は x の平方根を表す」)、6 5 6 0 個のベクトルを領域中心ベクトルとし、部分ベクトル v との内積が最も大きい領域中心ベクトル p_d を求めて、その番号 d を v の所属領域の領域番号とし、 p_j と v とのなす角の余弦を、偏角 c として求める。

偏角分布集計手段 2 0 6 は、領域番号算定手段 2 0 5 が 2 0 万個の各ベクトルデータの 3 7 個の部分ベクトルについて算定した偏角の値 c の分布を集計し、4 個の連続する実数範囲、

20 偏角区分 0 = $[c_0, c_1)$,

偏角区分 1 = $[c_1, c_2)$,

偏角区分 2 = $[c_2, c_3)$,

偏角区分 3 = $[c_3, c_4)$,

として偏角区分を定める。

25 偏角区分表 2 0 7 は、偏角分布集計手段 2 0 6 が算出した偏角区分を格納する。

ノルム区分番号算定手段 208 は、部分ベクトル算定手段 202 が算出した各部分ベクトルのノルムの属するノルム区分番号 r を、ノルム区分表 204 を検索して定める。

5 偏角区分番号算定手段 209 は、部分ベクトル算定手段 202 が算出した各部分ベクトル v と、領域番号算定手段 205 が v に対して算定した領域中心ベクトル p から、 v と p の偏角の属する偏角区分番号 c を、偏角区分表 207 を検索して定める。

索引データ算定手段 210 は、部分ベクトル算定手段 202 の算定した部分ベクトル v 、および部分空間番号 b と、領域番号算定手段 205
10 の算定した領域番号 d と、偏角区分番号算定手段 209 の算定した偏角区分番号 c と、ノルム区分番号算定手段 208 の算定したノルム区分番号 r とから検索用の鍵、

$$K = ((b * 6560 + d) * 4 + c) * 256 + r$$

を作成し、この鍵 K と部分ベクトルの識別番号 i と成分区分番号 y_j の
15 組 (K, i, y) を索引データとして算定する。

索引構成手段 211 は、索引データ算定手段 210 の算定した索引データ (K, i, y) から、 K を鍵とし、 (i, y) を検索するような探索木と、各識別番号 i と各部分空間番号 b の組に対して、領域番号 d と偏角区分番号 c と、ノルム区分番号 r とから第 2 の鍵、

20
$$L = (d * 4 + c) * 256 + r$$

を格納した逆探索表と、ノルム区分表 204 と、偏角区分表 207 とを格納した索引を構成する。

ベクトル索引 212 は、索引構成手段 211 が作成する探索木と逆探索表とノルム区分表 204 と偏角区分表 207 とを格納する。なお、構成
25 要素 201 ~ 212 は、それぞれ図 1 の構成要素 101 ~ 112 と対応し、特に構成要素 201 ~ 209 は、図 1 の構成要素 101 ~ 109

と同一である。

成分区分番号算定手段 2 1 3 は、部分ベクトル算定手段 2 0 2 が算定した部分ベクトル v_0 と、ノルム区分番号算定手段 2 0 8 が算定したノルム区分番号とから、部分ベクトルの各成分の値から、0 ~ 2 5 5 の範囲の成分区分番号 $y_0 \sim y_7$ を算定する。

(ベクトル索引作成装置の動作)

(ベクトル索引作成の第 1 ステップ)

以上の構成のベクトル索引作成装置において、その動作を図面をもとに説明する。ベクトル索引作成の第 1 ステップにおけるノルム区分表 R および偏角区分表 C の作成処理の手順は、第 1 の実施の形態における手順と同一であり、同一のベクトルデータベースならば、作成されるノルム区分表 R および偏角区分表 C は、ともに第 1 の実施の形態におけるノルム区分表 R および偏角区分表 C の内容と同一となるので、説明を省略する。

(ベクトル索引作成の第 2、第 3 ステップ)

図 7 A、7 B は一体となってベクトル索引作成の第 2、第 3 ステップにおける索引登録データの算定およびベクトル索引の作成処理の手順を記述した流れ図である。図 7 A、7 B のステップ 1 2 0 0 ~ ステップ 1 2 1 6 はそれぞれ図 6 A、6 B のステップ 1 1 0 0 ~ ステップ 1 1 1 6 に対応し、特に、ステップ 1 2 1 1、ステップ 1 2 1 5、ステップ 1 2 1 7 以外の各ステップは、図 6 A、6 B の対応するステップと同一の処理であるため、説明を省略する。

ステップ 1 2 1 7 では、部分ベクトル $u[0..7]$ から、 u の各成分に対応する成分区分番号 $y[0..7]$ を算定する。どの $u[m]$ に対しても、

$\text{abs}(u[m]) \leq |u| < R[r+1]$ なので、

$$-1 < u[m] / R[r+1] < +1$$

が成り立ち、成分区分番号 $y[m]$ は $0 \sim 255$ までの、8ビットで表現できる整数値となる。ステップ1211では、 u の代わりに、 y を使って、

- 5 (i, y) という、整数データを k を鍵にして探索木 T に登録する。各々の $y[m]$ は 8 ビットで表現できるので、 $u[m]$ を浮動小数点形式で登録する場合に比べ、探索木 T の容量が大幅に低減する。ステップ1215では、このようにして作成した探索木 T を含むベクトル索引を作成するので、結果的に作成したベクトル索引の容量が、 $u[m]$ を登録する場合に比べて小さくて済む。
- 10

なお、本実施の様態2においては、ステップ1217で8ビットの整数値 $y[m]$ で各成分 $u[m]$ を近似したが、8ビットでは類似検索時の精度が不足する場合には、十分な精度が得られるよう、9ビット～24ビットで表現して登録すればよい。

- 15 以上説明したように、本発明の第2の実施の形態のベクトル索引作成方法および装置によれば、以下のような優れた効果を奏する。

- 1) 296次元のベクトルを、37種の8次元の部分ベクトルに分解し、各々の部分ベクトルに対して、その方向を、6560種の領域中の所属領域の領域番号と偏角区分番号の組で精密に量子化し、その大きさをノルム区分番号で量子化し、さらに加えて部分ベクトルの各成分を成分区分番号という、ノルム区分を基準にした量子化を行って、複数の鍵を1つの整数値にエンコードして、近似結果である部分ベクトルの成分区分番号とともに探索木に登録することで、部分空間毎の高速で高精度な範囲検索が可能となる。
- 20

- 25 2) 逆探索表を作成・装備することで、ベクトルデータの識別番号を指定してベクトルの成分を取得する機能が、成分データを2重持ちする

ことなく実現できるため、元のベクトルデータベース 101 が検索時には不要となり、検索装置の記憶容量を削減することが可能となる。

3) ノルム分布集計手段、偏角分布集計手段において、各区分に所属する部分ベクトルの個数ができる限り均一になるように区分境界を定めるため、分布に偏りのあるベクトルデータベースに対しても、常に最適な（検索速度の低下を最小限に抑えた）ベクトル索引を作成することができる。

4) 領域中心ベクトルとして、成分が $\{0, +1, -1\}$ のいずれかである 0 ベクトルでない全てのベクトルを正規化したベクトルの集合を用いているため、各部分ベクトルの所属領域の算定が、領域の個数には依存せずに行え、部分ベクトルの成分の絶対値順序の算出と、成分の絶対値の加算などの、極めて少ない計算量で済むことから、数十万～数百万個のベクトルデータからなる大規模なベクトルデータベースの場合でも、実用的な処理時間でベクトル索引を作成することが可能となる。

5) 作成するベクトル索引の容量を大幅に削減することができる。

<第 3 の実施の形態>

次に、本発明の第 3 の形態について、図面を参照しながら説明する。

20 (類似ベクトル検索装置の構成)

図 3 は本発明の請求項 9、11、12、22、24、25 にかかる、類似ベクトル検索装置の全体構成を表わすブロック図である。図 3 において、ベクトル索引 301 は、上記した第 1 の実施の形態のベクトル索引作成装置によって作成したものであり、20 万件の新聞記事を集めた新聞記事全文データベースから作成した、各新聞記事の特徴を表す 296 次元の実ベクトルと、各記事を一意に識別するための 1 から 2000

00の範囲の識別番号の2項目からなるベクトルデータを20万個格納した、図12A、12Bのような内容を持つベクトルデータベースから作成したベクトル索引である。

5 検索条件入力手段302は、前記新聞記事全文データベースに対する類似検索を行うために、前記新聞記事全文データベース中のいずれかの記事の識別番号と、類似検索範囲を示す0～100までの類似度下限値と最大取得件数とを入力し、ベクトル索引301を識別番号で検索して入力された識別番号から対応する記事のベクトルを質問ベクトルQとして求め、類似度下限値から内積下限値 α を求める。

10 部分質問条件算定手段303は、検索条件入力手段302が求めた質問ベクトルQに対し、0～36までの部分空間について、37種の8次元部分質問ベクトルqと、qに対応する部分ベクトルとの内積の下限値である部分内積下限値fを $f = \alpha |q|^2 / |Q|^2$ によって算定する。

15 検索対象範囲生成手段304は、部分質問条件算定手段303が求めた部分空間bに対する部分質問ベクトルqと部分内積下限値fとベクトル索引301中のノルム区分表および偏角区分表とから、部分質問ベクトルqとの部分内積が部分内積下限値fより大きい可能性がある部分文書ベクトルを含んでいる領域を特定する領域番号d、偏角区分番号c、ノルム区分範囲 $[r_1, r_2]$ の組(d, c, $[r_1, r_2]$)を全て列挙する。

索引検索手段305は、検索対象範囲生成手段304が各部分空間bに対して生成した(d, c, $[r_1, r_2]$)から、ベクトル索引301に対する検索条件Kを、ベクトル索引作成時における鍵の算定と同様に、

$$K = [k_{min}, k_{max}]$$

25 $k_{min} = b * 7617440 + d * 1024 + c * 256 + r_1$

$$k_{max} = b * 7617440 + d * 1024 + c * 256 + r_2$$

と算定し、ベクトル索引 301 を検索条件 K で範囲検索して検索条件に合致する鍵を持つ部分ベクトル v と識別番号 i との組 (i, v) をすべて求める。

内積差分上限算定手段 306 は、索引検索手段 305 が求めた部分ベクトル v と識別番号 i との組 (i, v) および部分質問条件算定手段 303 が求めた部分質問ベクトル q と部分内積下限値 f から、 $t = (v \cdot q) - f$ によって部分内積差分値 t を算定して、識別番号 i を添字とする表の要素 $S[i]$ に、部分内積差分値 t を累算（加算）することで、識別番号 i のベクトルデータのベクトル V と質問ベクトル Q との内積 $Q \cdot V$ から内積下限値 α を減じた内積差分の上限値を算定する。

内積差分表 307 は、内積差分上限算定手段 306 が算出した内積差分の上限値を蓄積するものであり、識別番号 i のベクトルデータの内積差分値 $S[i]$ を参照／格納する。

類似検索決定手段 308 は、内積差分表 307 の要素 $S[i]$ 中で、内積差分上限値 $S[i]$ が正の値で大きいものから順に、ベクトル索引 301 を識別番号 i で検索して対応するベクトル V を求め、 V と検索条件入力手段 302 が算出した質問ベクトル Q との内積 $V \cdot Q$ から、検索条件入力手段 302 が算出した内積下限値 α を差し引いた内積差分値 $V \cdot Q - \alpha$ を算出して、 $S[i]$ を内積差分値 $V \cdot Q - \alpha$ で置き換え、内積差分値を算出していない記事の部分内積差分累算値の最大値より大きい内積差分値を持つ、内積差分値を算出した記事が L 個以上になった時点、あるいは部分内積差分累算値が正である全ての記事の内積差分値を算出した時点で、内積差分値が正である、内積差分値の大きな最大 L 個の結果候補 $(i, S[i])$ に対し、識別番号 i と内積 $S[i] + \alpha$ との組 $(i, S[i] + \alpha)$ を、検索結果として、検索結果出力手段 309 に出力する。

検索結果出力手段 309 は、類似ベクトル検索の結果である、最大 L

件の新聞記事の識別番号と0～100の範囲の類似度とを、類似結果決定手段308が求めた検索結果から算定して表示する。

(類似ベクトル検索装置の動作)

- 5 以上の構成の類似ベクトル検索装置において、その動作を図面をもとに説明する。図8A、8Bは一体となつては類似ベクトル検索の第1ステップにおける検索処理の手順を記述した流れ図、図9は類似ベクトル検索の第2ステップにおける検索処理の手順を記述した流れ図である。類似ベクトル検索の第1ステップでは、検索条件入力手段302から入力された検索条件から部分質問ベクトル q と部分内積下限値 f とを作成し、ベクトル索引301を検索して、各ベクトルデータの内積差分上限値 $S[i]$ 、即ち質問ベクトルとの内積から内積下限値を減じた値が $S[i]$ 未満となるような値を内積差分表307に求める。次に、類似ベクトル検索の第2ステップでは、第1ステップで内積差分表307に求めた内積差分上限値を手がかりにして、類似検索結果決定手段308が、
- 10 「質問ベクトルとの内積が α より大きい」という検索条件に合致し、かつ質問ベクトルとの内積が比較的大きいベクトルデータから順に、ベクトルの成分を検索して内積差分を求めていき、内積差分を求めていないベクトルデータのどれよりも内積差分値が大きいことが保証されるベクトルデータが指定件数（即ち L 件）以上集まるか、あるいは検索条件に
- 15 合致する全てのベクトルデータについて内積差分値を求めるまで処理を続け、得られた内積差分値から内積を算出して最終結果を出力する。

(類似ベクトル検索の第1ステップ)

- 25 以下、検索条件として識別番号1、類似度下限値90、最大取得件数10が入力された場合を例にとって、図8A、8B、図9にそつて類似

ベクトル検索の内容を説明する。識別番号が1なので、296次元のベクトルの各成分は、図12Aの通りとなる。まず、ステップ1301で内積差分表Sの20万個の要素S[0]～S[200000]を0に初期設定する。次に、上記の検索条件を検索条件入力手段302から読み込み、それぞれi、Z、Lに格納する。

ステップ1302で部分空間番号bを0に初期化した後、類似度下限値Zから、内積下限値 α を算出する。この検索条件の場合には、 $\alpha \leftarrow (90 - 50) / 50 = 0.8$ となる。ステップ1304、ステップ1305で、各部分空間について、ベクトル索引301の逆変換表Kを引いて鍵を求め、探索表を検索してベクトルデータを求め、識別番号が1であるもののベクトル部分をQに格納することで、質問ベクトルをQ[0..295]に求める。ステップ1306で部分空間番号を初期化した後、ステップ1307～ステップ1317で各部分空間についてベクトル索引を検索して内積差分表307に各ベクトルデータの内積差分上限値を求める。

ステップ1307では、部分空間番号bの部分質問ベクトルq[0..7]と部分内積下限値f、即ちこの部分空間の部分ベクトルデータとqとの内積の下限値を求める。b=0の場合には、 $|q|^2 = 0.221795$ 、 $|Q|^2 = 1$ であるから、

20 $f = 0.8 * 0.221795 / 1.0 = 0.177436$

となる。領域番号dを0に初期化したあと、検索対象範囲を決定するために用いる表Wを作成する。表Wは偏角区分番号c、ノルム区分番号rで引いて、着目している領域番号dの領域の中心ベクトルpと部分質問ベクトルqとの内積 $p \cdot q$ がW[c, r]未満ならば、区分(d, c, 0)～区分(d, c, r)の部分ベクトルvと部分質問ベクトルqとの内積がf以下となるように作成する。この場合には、区分(d, c, 0)

～区分 (d, c, r) の部分ベクトルはこの部分空間に対する検索条件 (即ち部分内積が f より大きい) を満たさないため、これらの区分の検索を省略することができる。

表 W を求めるには、部分ベクトル v が領域 d 内で最も部分質問ベクトル q に近い場合、p, q, v が一平面上にあり、かつ v と q とのなす角 ω が偏角区分 c の範囲内でもっとも小さい場合を考えればよい。このとき、p と q のなす角を θ 、p と v のなす角の最大値を ϕ とすると、v と q のなす角 ω は $\omega = \theta - \phi$ となるから、

$$f < v \cdot q = |v| * |q| * \cos(\theta - \phi) < R[r+1] * |q| * (\cos \theta * \cos \phi + \sin \theta \sin \phi)$$

と、

$$C[c] = \cos \phi$$

$$\cos \theta = (p \cdot q) / |p| * |q| = (p \cdot q) / |q|$$

とから、p · q が満足する不等式、

$$f < R[r+1] * C[c] * (p \cdot q) + R[r+1] * \sqrt{1 - C[c]^2} * \sqrt{(|q|^2 - (p \cdot q)^2)}$$

を解いて、ステップ 1307 の W [c, r] の式を得る。

このように、実際の部分ベクトル v の成分を参照せず、また領域 d にも依存せずに、部分質問ベクトルのノルム |q| だけから、表 W [c, r] の値を定めることができる。本実施の形態の場合には、ノルム区分表 R、偏角区分表 C がそれぞれ図 15 A、15 B、16 の通りなので、b = 0 の場合には、表 W は図 17 A、17 B のような内容となる。図中で表の値が「9. 9 9 9 9」となっている要素は、この部分質問ベクトル q については、ノルムが小さすぎて、どの方向の部分ベクトル v でも、q との内積が f に達することができないため、検索対象とはならないノルム区分であることを意味している。図 17 A、17 B からは、c = 0、

即ち偏角の値が大きい場合には、広範囲な検索を行い、 $c = 3$ 、即ち偏角の値が小さい場合には、ノルムの大きな部分のみの、より狭い範囲を検索するような内容になっていることがわかる。

ステップ 1 3 0 8 では、着目する領域の中心ベクトル p と部分質問ベ
 5 クトル q との内積 t を求め、偏角区分用のループ変数 c を 0 に初期化し
 た後、ステップ 1 3 0 9 で表 W の最小値を与える要素 $W[0, 255]$ より内
 積 t が小さいか否か検査し、小さい場合には、領域 d を鍵の一部とする
 どの部分ベクトルも検索条件を満足しないことが確定するのでステップ
 1 3 1 2 に飛び、そうでない場合にはステップ 1 3 1 0 で偏角区分 c に
 10 ついて、検索すべきノルム区分の最低値 r をステップ 1 3 0 7 で算定し
 た表 W を手がかりにして求め、この r と部分空間番号 b 、領域番号 d 、
 偏角区分番号 c とから、ベクトル索引 3 0 1 の検索範囲 $[k_{\min}, k_{\max}]$
 を求める。ステップ 1 3 1 1 ではこの検索範囲 $[k_{\min}, k_{\max}]$ を鍵にし
 て探索木を範囲検索し、範囲検索結果に含まれる、識別番号 j とベクト
 15 ル v の組 (j, v) のそれぞれについて、部分質問ベクトル q と v との
 内積から部分内積下限値 f を減じた部分内積差分値を算定、内積差分表
 3 0 7 の該当要素 $S[j]$ に累算する。

例えば、 $b = 0$ 、 $d = 4 2 1 2$ の場合には、

$q = (+0.029259 \ -0.016005 \ -0.021118 \ +0.024992 \ -0.006860$
 20 $-0.009032 \ -0.007255 \ -0.007715)$

$p_0 = (+1/2, -1/2, -1/2, +1/2, 0, 0, 0, 0)$
 であるから、

$$t = p \cdot q = +0.045687$$

となり、 $W[0, 255] = -0.02527$ よりも t が大きく、ステップ 1 3 1 0 に
 25 進む。図 1 7 A、1 7 B の表 W より、

$$W[0, r] \leq t < W[0, r + 1]$$

となるノルム区分番号 r は $r = 1$ であり、 $c = 0$ の場合には、探索木の鍵は、

$$[k_{\min}, k_{\max}] = [0*6717440+4212*1024+0*256+1, 0*6717440+4212*1024+0*256+255]$$

$$5 \quad = [4313089, 4313343]$$

となる。識別番号 1 のベクトルデータの $b = 0$ の部分ベクトル

$v = (+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.007255 -0.007715)$ は、鍵 $k = 0*6717440+4212*1024+0*256+1 = 4313089$ で登録されているので、この範囲検索の結果の 1 つとなる。そ

10 の部分内積差分値は、

$$(v \cdot q) - f = 0.221795 - 0.177436 = 0.044359$$

となり、 $S[1] = 0.044359$ となる。

また、識別番号 2 のベクトルデータの $b = 0$ の部分ベクトル

$v = (+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.007255 -0.007715)$ は、鍵 $k = 0*6717440+619*1024+2*256+2$ で登録されており、 $b = 0$ 、 $c = 2$ 、 $d = 619$ の範囲検索の結果に含まれ、その部分内積差分値は、

$$(v \cdot q) - f = 0.00005$$

となり、 $S[2] = 0.00005$ となる。

20 同様に、 $b = 1$ では、識別番号 2 のベクトルデータの部分ベクトルは鍵 $k = 1*6717440+2691*1024+1*256+93$ で登録されており、 $b = 1$ 、 $c = 1$ 、 $d = 2691$ の範囲検索の結果に含まれ、その部分内積差分値は、

$$(v \cdot q) - f = 0.00217$$

が $S[2]$ に累算され、 $S[2] = 0.00222$ となる。

25 このようにして、ステップ 1312、ステップ 1313 で c を増やしながら各偏角区分について検索範囲の決定と検索処理、内積差分の算出

- と累算を行う。次いで、ステップ1314、ステップ1315で領域番号dを6560まで順に増やしながら各領域についてステップ1308～ステップ1313の処理を行う。さらにステップ1316、ステップ1317で部分空間番号を37まで順に増やしながら各部分空間について
- 5 ステップ1307～ステップ1315の処理を行い、類似ベクトル検索の第1ステップを終了する。この段階で、内積差分表307には各識別番号のベクトルデータVについて、質問ベクトルQとの内積 $V \cdot Q$ と、内積下限値 α との差である、内積差分値 $(V \cdot Q) - \alpha$ の推定値の上限が求まる。なぜなら、各部分空間bについて、部分質問ベクトルqとの
- 10 内積が部分内積下限値fより大きな部分ベクトルについては、もれなく部分内積差分値を求めているので、部分内積差分値を求めなかったベクトルデータの部分内積差分値は負の値となるはずであり、この負の値を0で置き換えて累算している（「内積差分表を変化させない」ということは0を累算していることに等価である）ので、部分内積差分値の累算
- 15 結果は、内積差分値を上から押さえる、内積差分上限値の1つとなるからである。以上のようにして内積差分表307を求めた後、類似ベクトル検索の第2ステップを実行し、最終的な検索結果を得る。

（類似ベクトル検索の第2ステップ）

- 20 次に、図9の流れ図にそって第2ステップの処理手順を説明する。ステップ1401で現時点での検索条件を満たす候補の件数を0クリアし、ベクトルデータの内積差分を求めたか否かのフラグA[0..200000]も0、即ち「内積差分を求めている」に初期設定する。そして、現時点で検索条件を満たす候補のうちの内積差分値の最低値（＝しきい値）tを0
- 25 に初期化する。

ステップ1402では、まだ未調査、即ち内積差分を求めているべ

クトルデータがあるかどうか調べ、全てのベクトルデータの内積差分を求めた場合にはステップ1412に飛ぶ。ただし、検索条件として与えた内積下限値が0以上で、かつ、ベクトルデータの各成分の分布に偏りが小さい場合には、全てのベクトルデータの内積差分を求めるはるか前にステップ1404で条件が「no」になるため、通常の実行条件ではステップ1402が「no」になることはない。

ステップ1403では $A[j]$ が0、即ち未調査のベクトルデータ中で、内積差分表の値 $S[j]$ が最大となるようなベクトルデータの識別番号 j を求める。このステップの処理は、内積差分表307を、内積差分値の降順に整列するか、あるいはヒープなどのデータ構造で表現することにより、効率のよい実行が可能である。

ステップ1404では、先に求めた t と $S[j]$ とを比較し、もしも $S[j]$ が t 以下であれば、未調査のベクトルデータ中には現時点の n 件の候補の内積差分値を超えるベクトルデータは存在しないことが確定するのでステップ1412に飛び、現時点での候補から結果を計算して出力し、検索処理を終了する。 t が $S[j]$ より大きい場合には、ステップ1405で着目ベクトルデータのフラグ $A[j]$ を1に変えて「内積差分を求めた」ことを記録し、ベクトル索引301を検索して識別番号 j のベクトル V を求め、質問ベクトル Q との内積差分値 $(V \cdot Q) - \alpha$ を求めて内積差分表207の該当要素 $S[j]$ 内の上限値を、正確な内積差分値で置き換える。記憶領域に余裕のある場合には、内積差分表を置き換えず、新たな表に記録しても差し支えない。

ステップ1406では、置き換えた後の $S[j]$ と t とを再び比較し、 $S[j]$ が t より大きい場合には、ステップ1407～ステップ1411までを実行して識別番号 j のベクトルデータを候補に加える。ステップ1407では、この時点で既に L 件の候補が求まっているかどうか判

定し、まだ L 件求まっていない場合には、ステップ1408で候補の件数 n を増やす。ステップ1409では候補の識別番号の配列 B の最終候補（内積差分が候補中でもっとも低いような候補）として j を登録した後、 $B[0..n-1]$ を、 $S[B[k]]$ の降順に整列する。ステップ1410で

5 候補件数 n が L に達していれば、ステップ1411でしきい値 t を更新し、ステップ1402に戻って処理を続ける。

ステップ1402、もしくはステップ1404で判定が「no」になると、上記のループを抜け、ステップ1412に進む。ステップ1412では、 n 件（最高 L 件）の候補の識別番号 $B[0] \sim B[n-1]$ のそれぞれ

10 について、既に求めた内積差分値 $S[B[k]]$ に α を加えて内積の値を求め、0から $n-1$ までの各 k について、内積が k 番目に大きいベクトルデータの結果番号 $B[k]$ と、質問ベクトル V との内積の値 $S[B[k]]$ との組 $(B[k], S[B[k]])$ を類似ベクトル検索の最終結果として出力し、類似ベクトル検索を終了する。

15 検索条件中の内積下限の値が0.5以上と十分に大きく、かつベクトルデータの分布に大きな偏りがなく、内積下限 α 以上の内積となるベクトルデータの個数が取得件数 L よりも十分多い場合には、ステップ1402～ステップ1411のループを、取得件数 L の数倍程度繰り返すと、ステップ1404の判定が「no」となるため、実際にベクトルを検索

20 して内積を求めるベクトルデータの個数が非常に少なく、最終結果を効率よく求めることが可能となる。しかも、この特性は L が数百程度の場合でも成り立つため、 L が比較的大きい検索条件の場合には、 L が高々数個の場合にしか実用的な検索速度が得られない従来の類似ベクトル検索方法に比べ、処理効率が非常によくなる。

25 以上説明したように、本発明の第3の実施の形態の類似ベクトル検索方法および装置によれば、ベクトルの次元が数百に及ぶベクトルデータ

- を多数集めたベクトルデータベースに対して、「最も類似している L 個のベクトルデータを求める」という型の高速な類似検索が可能で、さらに L が比較的大きい（数十～数百個）場合でも検索処理が極端に遅くならず、「内積の値が 0.8 以上」といった類似検索範囲の指定が可能な、
- 5 ベクトルの内積を類似尺度として用いる優れた類似ベクトル検索方法および装置を提供することができる。

- なお、本実施の形態 3 においては、本発明の実施の形態 1 のベクトル索引作成装置で作成したベクトル索引を検索する場合について説明したが、実施の形態 1 の索引作成装置において、各部分ベクトルを求める処理を、
- 10 ノルム区分番号と各成分の成分区分番号とから各成分の値を求めるように変更するだけで、本実施の形態 3 の類似ベクトル検索装置は、実施の形態 2 のベクトル索引作成装置で作成したベクトル索引の検索にも用いることができ、上記と同様な効果が期待できる。

- さらに、本実施の形態 3 においては、類似ベクトル検索の第 1 ステップにおいて、各部分空間 b についての検索処理を逐次的に行う手順を説明したが、図 $8A$ 、 $8B$ の流れ図のステップ 1306 ～ステップ 1317 のループは、多数のCPU（中央処理装置）を持つ並列計算機の場合には、それぞれのCPUで分担して処理し、共通の内積差分表に中間結果として累算することで、容易に高い並列度で並列処理することができ、
- 20 検索速度をさらに向上させることが可能である。

<第4の実施の形態>

次に、本発明の第4の形態について、図面を参照しながら説明する。

- 25 （類似ベクトル検索装置の構成）

図4は本発明の請求項 10 、 11 、 13 、 23 、 24 、 26 にかかる、

類似ベクトル検索装置の全体構成を表わすブロック図である。図4において、ベクトル索引401は、上記した第1の実施の形態のベクトル索引作成装置によって作成したものであり、20万件の新聞記事を集めた新聞記事全文データベースから作成した、各新聞記事の特徴を表す29
5 6次元の実ベクトルと、各記事を一意に識別するための1から200000の範囲の識別番号の2項目からなるベクトルデータを20万個格納した、図12A、12Bのような内容を持つベクトルデータベースから作成したベクトル索引である。

検索条件入力手段402は、前記新聞記事全文データベースに対する
10 類似検索を行うために、前記新聞記事全文データベース中のいずれかの記事の識別番号と、類似検索範囲を示す0～100までの類似度下限値と最大取得件数とを入力し、ベクトル索引401を識別番号で検索して入力された識別番号から対応する記事のベクトルを質問ベクトルQとして求め、類似度下限値から二乗距離、即ち距離の二乗の上限値である、
15 二乗距離上限値 α^2 を求める。

部分質問条件算定手段403は、検索条件入力手段402が求めた質問ベクトルQに対し、0～36までの部分空間について、37種の8次元部分質問ベクトルqと、qに対応する部分ベクトルとの二乗距離の上限値である部分二乗距離上限値 f^2 を $f^2 = \alpha^2 |q|^2 / |Q|^2$ によって算定する。
20

検索対象範囲生成手段404は、部分質問条件算定手段403が求めた部分空間bに対する部分質問ベクトルqと部分二乗距離上限値 f^2 とベクトル索引401中のノルム区分表および偏角区分表とから、部分質問ベクトルqとの部分二乗距離が部分二乗距離上限値 f^2 より小さい可能性のある部分ベクトルを含んでいる領域を特定する領域番号d、偏角区分番号c、ノルム区分範囲 $[r_1, r_2]$ の組(d, c, $[r_1, r_2]$)
25

を全て列挙する。

索引検索手段 405 は、検索対象範囲生成手段 404 が各部分空間 b に対して生成した $(d, c, [r_1, r_2])$ から、ベクトル索引 401 に対する検索条件 K を、ベクトル索引作成時における鍵の算定と同様に、

$$5 \quad K = [k_{min}, k_{max}]$$

$$k_{min} = b * 7617440 + d * 1024 + c * 256 + r_1$$

$$k_{max} = b * 7617440 + d * 1024 + c * 256 + r_2$$

と算定し、ベクトル索引 401 を検索条件 K で範囲検索して検索条件に合致する鍵を持つ部分ベクトル v と識別番号 i との組 (i, v) をすべ

10 て求める。

二乗距離差分上限算定手段 406 は、索引検索手段 405 が求めた部分ベクトル v と識別番号 i との組 (i, v) および部分質問条件算定手段 403 が求めた部分質問ベクトル q と部分二乗距離上限値 f^2 から、

$$t = f^2 - |v - q|^2$$

15 番号 i を添字とする表の要素 $S[i]$ に、部分二乗距離差分値 t を累算

(加算) することで、二乗距離上限値 α^2 から、識別番号 i のベクトルデータのベクトル V と質問ベクトル Q との二乗距離 $|V - Q|^2$ を減じた二乗距離差分の上限値を算定する。

20 二乗距離差分表 407 は、二乗距離差分上限算定手段 406 が二乗距離差分の上限値を蓄積する、識別番号 i のベクトルベータの二乗距離差分値 $S[i]$ を参照／格納する。

類似検索結果決定手段 408 は、二乗距離差分表 407 の要素 $S[i]$ 中で、二乗距離差分上限値 $S[i]$ が正の値で大きいものから順に、ベクトル索引 401 を識別番号 i で検索して対応するベクトル V を求め、検索条件入力手段 402 が算出した二乗距離上限値 α^2 から、 V と検索条件入力手段 402 が算出した質問ベクトル Q との二乗距離 $|V - Q|^2$ を

- 差し引いた二乗距離差分値 $\alpha^2 - |V - Q|^2$ を算出して、 $S[i]$ を二乗距離差分値 $\alpha^2 - |V - Q|^2$ で置き換え、二乗距離差分値を算出していない記事の部分二乗距離差分累算値の最大値より大きい二乗距離差分値を持った二乗距離差分値を算出した記事が L 個以上になった時点、あるいは部分二乗距離差分累算値が正である全ての記事の二乗距離差分値を算出した時点で、二乗距離差分値が正である、二乗距離差分値の大きな最大 L 個の結果候補 $(i, S[i])$ に対し、識別番号 i と距離 $sqr t(\alpha^2 - S[i])$ との組 $(i, sqr t(\alpha^2 - S[i]))$ を、検索結果として、検索結果出力手段に出力する。
- 10 検索結果出力手段 409 は、類似ベクトル検索の結果である、最大 L 件の新聞記事の識別番号と $0 \sim 100$ の範囲の類似度とを、類似検索結果決定手段 408 が求めた検索結果から算定して表示するものである。

(類似ベクトル検索装置の動作)

- 15 以上の構成の類似ベクトル検索装置において、その動作を図面をもとに説明する。図 10A、10B は一体となって類似ベクトル検索の第 1 ステップにおける検索処理の手順を記述した流れ図、図 11A、11B は一体となって類似ベクトル検索の第 2 ステップにおける検索処理の手順を記述した流れ図である。類似ベクトル検索の第 1 ステップでは、検索条件入力手段 402 から入力された検索条件から部分質問ベクトル q と部分二乗距離上限値 f とを作成し、ベクトル索引 401 を検索して、各ベクトルデータの二乗距離差分上限値 $S[i]$ 、即ち二乗距離上限値から質問ベクトルとの二乗距離を減じた値が $S[i]$ 未満となるような値を二乗距離差分表 407 に求める。次に、類似ベクトル検索の第 2 ステップでは、第 1 ステップで二乗距離差分表 407 に求めた二乗距離差分上限値を手がかりにして、類似検索結果決定手段 408 が、「質問ベ
- 20
- 25

クトルとの二乗距離が α^2 より小さい」という検索条件に合致し、かつ質問ベクトルとの二乗距離が比較的小さいベクトルデータから順に、ベクトルの成分を検索して二乗距離差分を求めていき、二乗距離差分を求めていないベクトルデータのどれよりも二乗距離差分値が大きいことが

5 保証されるベクトルデータが指定件数（即ちL件）以上集まるか、あるいは検索条件に合致する全てのベクトルデータについて二乗距離差分値を求めるまで処理を続け、得られた二乗距離差分値から距離を算出して最終結果を出力する。

10 （類似ベクトル検索の第1ステップ）

以下、検索条件として識別番号1、類似度下限値90、最大取得件数10が入力された場合を例にとって、図10A、10B、11A、11Bにそって類似ベクトル検索の内容を説明する。識別番号が1なので、296次元のベクトルの各成分は、図12Aの通りとなる。まず、ステップ1501で二乗距離差分表Sの20万個の要素S[0]～S[200000]

15 [0000]を0に初期設定する。次に、上記の検索条件を検索条件入力手段402から読み込み、それぞれi、Z、Lに格納する。

ステップ1502で部分空間番号bを0に初期化したあと、類似度下限値Zから、二乗距離上限値 α^2 を算出する。この検索条件の場合には、

20 $\alpha \leftarrow (100 - 90) / 50 = 0.2$ となる。ステップ1504、ステップ1505で、各部分空間について、ベクトル索引401の逆変換表Kを引いて鍵を求め、探索表を検索してベクトルデータを求め、識別番号が1であるもののベクトル部分をQに格納することで、質問ベクトルをQ[0..295]に求める。ステップ1506で部分空間番号を初期化した後、ステップ1507～ステップ1517で各部分空間についてベク

25 トル索引を検索して二乗距離差分表407に各ベクトルデータの二乗距

離差分上限値を求める。

ステップ1507では、部分空間番号bの部分質問ベクトルq [0..7]と部分二乗距離上限値 f^2 、即ちこの部分空間の部分ベクトルデータvとqとの部分二乗距離の上限値を求める。b=0の場合には、 $|q|^2 = 0.221795$ 、 $|Q|^2 = 1$ であるから、

$$f^2 = 0.04 * 0.221795 / 1.0 = 0.0088718$$

となる。領域番号dを0に初期化したあと、検索対象範囲を決定するために用いる表Wを作成する。表Wは偏角区分番号c、ノルム区分番号rで引いて、着目している領域番号dの領域の中心ベクトルpと部分質問ベクトルqとの内積 $p \cdot q$ がW[c, r]未満ならば、区分(d, c, 0)～区分(d, c, r)の部分ベクトルvと部分質問ベクトルqとの部分二乗距離が f^2 以上となるように作成する。この場合には、区分(d, c, 0)～区分(d, c, r)の部分ベクトルはこの部分空間に対する検索条件(即ち部分二乗距離が f^2 より小さい)を満たさないため、これらの区分の検索を省略することができる。

表Wを求めるには、部分ベクトルvが領域d内で最も部分質問ベクトルqに近い場合、即ちp、q、vが一平面上にあり、かつvとqとのなす角 ω が偏角区分cの範囲内でもっとも小さい場合を考えればよい。このとき、pとqのなす角を θ 、pとvのなす角の最大値を ϕ とすると、vとqのなす角 ω は $\omega = \theta - \phi$ となるから、

$$f^2 > |v - q|^2 = |v|^2 + |q|^2 - 2 * |v| * |q| * \cos(\theta - \phi) > R[r]^2 + |q|^2 - 2 * R[r+1] * |q| * (\cos \theta * \cos \phi + \sin \theta \sin \phi)$$

と、

$$C[c] = \cos \phi$$

$$\cos \theta = (p \cdot q) / |p| * |q| = (p \cdot q) / |q|$$

とから、 $p \cdot q$ が満足する不等式

$$f^2 > R[r]^2 + |q|^2 - 2 * R[r+1] * ((p \cdot q) * C[c] + \sqrt{(|q|^2 - (p \cdot q)^2) * \sqrt{1 - C[c]^2}})$$

を解いて、ステップ1507の $W[c, r]$ の式を得る。

- 5 このように、実際の部分ベクトル v の成分を参照せず、また領域 d にも依存せずに、部分質問ベクトルのノルム $|q|$ だけから表 $W[c, r]$ の値を定めることができる。本実施例の場合には、ノルム区分表 R 、偏角区分表 C がそれぞれ図15A、15B、16の通りなので、 $b=0$ 、 $b=1$ の場合には、表 W は図18A、18B、18Cのような内容となる。
- 10 図17A、17Bと同様に、図中で表の値が「9.99999」となっている要素は、この部分質問ベクトル q については検索対象とはならないノルム区分であることを意味している。また、 $b=0$ では区分10～255、 $b=1$ 区分0～59および区分180～255までの表の値が記載されていないが、この部分はすべて値が「9.99999」である
- 15 ので省略している。これが、類似尺度として距離を用いているため、ノルムが小さすぎても、逆に大きすぎても、部分質問ベクトルとの距離が大きくなってしまい、結果的に「距離が α 未満」という検索条件を満足できなくなるためである。

- ステップ1508で着目する領域の領域中心ベクトル p と部分質問ベ
- 20 クトル q との内積 t を求め、偏角区分用のループ変数 c を0に初期化した後、ステップ1509で表 W の最小値を与える要素 $\text{Min}(W[0, r])$ より内積 t が小さいか否か検査し、小さい場合には、領域 d を鍵の一部とするどの部分ベクトルも検索条件を満足しないことが確定するのでステップ1512に飛び、そうでない場合には、ステップ1510で偏角
- 25 区分 c について、検索すべきノルム区分の最低値 r_{\min} 及び最高値 r_{\max} をステップ1507で算定した表 W を手がかりにして、 $W[c, r]$

t が成り立つノルム区分番号 r の区間として求め、この $[r_{\min}, r_{\max}]$ と部分空間番号 b 、領域番号 d 、偏角区分番号 c とから、ベクトル索引 401 の検索範囲 $[k_{\min}, k_{\max}]$ を求める。

ステップ 1511 では、この検索範囲 $[k_{\min}, k_{\max}]$ を鍵にして探索木を範囲検索し、範囲検索結果に含まれる、識別番号 j とベクトル v の組 (j, v) のそれぞれについて、部分二乗距離上限値 f^2 から部分質問ベクトル q と v との部分二乗距離 $|v - q|^2$ を減じた部分二乗距離差分値を算定、二乗距離差分表 407 の該当要素 $S[j]$ に累算する。例えば、 $b = 0$ 、 $d = 4212$ の場合には、

$$10 \quad q = (+0.029259 \ -0.016005 \ -0.021118 \ +0.024992 \ -0.006860 \\ -0.009032 \ -0.007255 \ -0.007715)$$

$$p = (+1/2, -1/2, -1/2, +1/2, 0, 0, 0, 0)$$

であるから、

$$t = p \cdot q = +0.045687$$

15 となり、 $\text{Min}(W[0, r]) = 0.03356$ よりも t が大きく、ステップ 1510 に進む。図 15A、15B の表 W より、例えば $c = 0$ の場合には、

$$r_{\min} = 1, \quad r_{\max} = 5$$

であり、探索木の検索範囲は、

$$20 \quad [k_{\min}, k_{\max}] = [0*6717440+4212*1024+0*256+1, \\ 0*6717440+4212*1024+0*256+5] = [4313089, 4313093]$$

となる。識別番号 1 のベクトルデータの $b = 0$ の部分ベクトル x は、

$$x = (+0.029259 \ -0.016005 \ -0.021118 \ +0.024992 \ -0.006860 \ -0.009032 \\ -0.007255 \ -0.007715)$$

となり、 $k = 0*6717440+4212*1024+0*256+1 = 4313089$ で登録されている
25 ので、この範囲検索の結果の 1 つとなる。その部分二乗距離差分値は、

$$f^2 - |v - q|^2 = 0.0088718 - 0 = 0.0088718$$

となり、 $S[i] = 0.0088718$ となる。

- このようにして、ステップ1512、ステップ1513で c を増やしながら各偏角区分について検索範囲の決定と検索処理、二乗距離差分の
- 5 算出と累算を行う。ステップ1514、ステップ1515で領域番号 d を6560まで順に増やしながら各領域についてステップ1508～ステップ1513の処理を行う。ステップ1516、ステップ1517で部分空間番号を37まで順に増やしながら各部分空間についてステップ1507～ステップ1515の処理を行い、類似ベクトル検索の第1ス
- 10 テップを終了する。この段階で、二乗距離差分表407には各識別番号のベクトルデータ V について、二乗距離上限値 α^2 と、質問ベクトル Q との二乗距離 $|V - Q|^2$ との差である、二乗距離差分値 $\alpha^2 - |V - Q|^2$ の推定値の上限が求まる。なぜなら、各部分空間 b について、部分質問ベクトル q との二乗距離が部分二乗距離上限値 f^2 より小さな部分
- 15 ベクトルについては、もれなく部分二乗距離差分値を求めているので、部分二乗距離差分値を求めなかったベクトルデータの部分二乗距離差分値は負の値となるはずであり、この負の値を0で置き換えて累算している（「二乗距離差分表を変化させない」ということは0を累算していることに等価である）ので、部分二乗距離差分値の累算結果は、二乗距離
- 20 差分値を上から押さえる、二乗距離差分上限値の1つとなるからである。こうして二乗距離差分表407を求めた後、類似ベクトル検索の第2ステップを実行し、最終的な検索結果を得る。

（類似ベクトル検索の第2ステップ）

- 25 次に、図11A、11Bの流れ図にそって第2ステップの処理手順を説明する。ステップ1601で現時点での検索条件を満たす候補の件数

を0クリアし、ベクトルデータの二乗距離差分を求めたか否かのフラグ $A[0..200000]$ も0、即ち「二乗距離差分を求めている」に初期設定する。そして、現時点で検索条件を満たす候補のうちの二乗距離差分値の最低値(=しきい値) t を0に初期化する。

- 5 ステップ1602では、まだ未調査、即ち二乗距離差分を求めているベクトルデータがあるかどうか調べ、全てのベクトルデータの二乗距離差分を求めた場合にはステップ1612に飛ぶ。ただし、検索条件として与えた二乗距離上限値が1以下で、かつ、ベクトルデータの各成分の分布に偏りが小さい場合には、全てのベクトルデータの二乗距離差分
- 10 を求めるはるか前にステップ1604で条件が「no」になるため、通常の検索条件ではステップ1602が「no」になることはない。ステップ1603では $A[j]$ が0、即ち未調査のベクトルデータ中で、二乗距離差分表の値 $S[j]$ が最大となるようなベクトルデータの識別番号 j を求める。このステップの処理は、二乗距離差分表407を、二乗
- 15 距離差分値の降順に整列するか、あるいはヒープなどのデータ構造で表現することにより、効率のよい実行が可能である。

- ステップ1604では、先に求めた t と $S[j]$ とを比較し、もしも $S[j]$ が t 以下であれば、未調査のベクトルデータ中には現時点の n 件の候補の二乗距離差分値を超えるベクトルデータは存在しないことが
- 20 確定するのでステップ1612に飛び、現時点での候補から結果を計算して出力し、検索処理を終了する。

- t が $S[j]$ より大きい場合には、ステップ1605で着目ベクトルデータのフラグ $A[j]$ を1に変えて「二乗距離差分を求めた」ことを記録し、ベクトル索引401を検索して識別番号 j のベクトル V を求め、
- 25 質問ベクトル Q との二乗距離差分値 $\alpha^2 - |V - Q|^2$ を求めて二乗距離差分表407の該当要素 $S[j]$ 内の上限値を、正確な二乗距離差分

値で置き換える。記憶領域に余裕のある場合には、二乗距離差分表を置き換えず、新たな表に記録しても差し支えない。ステップ1606で、置き換えた後の $S[j]$ と t とを再び比較し、 $S[j]$ が t より大きい場合には、ステップ1607～ステップ1611までを実行して識別番号

5 号 j のベクトルデータを候補に加える。

ステップ1607では、この時点で既に L 件の候補が求まっているかどうか判定し、まだ L 件求まっていない場合には、ステップ1608で候補の件数 n を増やす。ステップ1609では候補の識別番号の配列 B の最終候補（二乗距離差分が候補中でもっとも低いような候補）として

10 j を登録した後、 $B[0..n-1]$ を、 $S[B[k]]$ の降順に整列する。ステップ1610で候補件数 n が L に達していれば、ステップ1611でしきい値 t を更新し、ステップ1602に戻って処理を続ける。ステップ1602、もしくはステップ1604で判定が「no」になると、上記のループを抜け、ステップ1612に進む。

15 ステップ1612では、 n 件（最高 L 件）の候補の識別番号 $B[0] \sim B[n-1]$ のそれぞれについて、既に求めた二乗距離差分値 $S[B[k]]$ から、 $sqr t(\alpha^2 - S[B[k]])$ で質問ベクトル Q との距離を求め、0から $n-1$ までの各 k について、距離が k 番目に小さいベクトルデータの結果番号 $B[k]$ と、質問ベクトル Q との距離の値 $S[B[k]]$ との組（ $B[k]$ 、

20 $S[B[k]]$ ）を類似ベクトル検索の最終結果として出力し、類似ベクトル検索を終了する。

検索条件中の二乗距離上限 α^2 の値が0.5以下と十分に小さく、かつベクトルデータの分布に大きな偏りがなく、二乗距離上限 α^2 未満の二乗距離となるベクトルデータの個数が取得件数 L よりも十分多い場合には、ステップ1602～ステップ1611のループを、取得件数 L の数倍程度繰り返すと、ステップ1604の判定が「no」となるため、

25

実際にベクトルを検索して二乗距離を求めるベクトルデータの個数が非常に少なく、最終結果を効率よく求めることが可能となる。しかも、この特性は L が数百程度の場合でも成り立つため、 L が比較的大きい検索条件の場合には、 L が高々数個の場合にしか実用的な検索速度が得られない従来の類似ベクトル検索方法に比べ、処理効率が非常によくなる。

以上、説明したように、本発明の第4の実施の形態の類似ベクトル検索方法によれば、ベクトルの次元が数百に及ぶベクトルデータを多数集めたベクトルデータベースに対して、「最も類似している L 個のベクトルデータを求める」という型の高速な類似検索が可能で、さらに L が比較的大きい(数十～数百個)場合でも検索処理が極端に遅くならず、「距離の値が0.2以下」といった類似検索範囲の指定が可能な、ベクトル間の距離を類似尺度として用いる優れた類似ベクトル検索方法が実現できる。

なお、本実施の形態4においては、本発明の実施の形態1のベクトル索引作成装置で作成したベクトル索引を検索する場合について説明したが、実施の形態1の索引作成装置において、各部分ベクトルを求める処理を、ノルム区分番号と各成分の成分区分番号とから各成分の値を求めるように変更するだけで、本実施の形態4の類似ベクトル検索装置は、実施の形態2のベクトル索引作成装置で作成したベクトル索引の検索にも用いることができ、上記と同様な効果が期待できる。

また、本実施の形態4においては、質問ベクトルを直接入力せず、ベクトルデータベース中のベクトルデータの識別番号を指定する形態を説明したが、外部から直接質問ベクトルのデータを指定した場合でも、上記の説明と同様の方法で、類似ベクトル検索装置を用意に実施できる。

さらに、本実施の形態4においては、類似ベクトル検索の第1ステップにおいて、各部分空間 b についての検索処理を逐次的に行う手順を説

明したが、図10A、10Bの流れ図のステップ1506～ステップ1517のループは、多数のCPU（中央処理装置）を持つ並列計算機の場合には、それぞれのCPUで分担して処理し、共通の内積差分表に中間結果として累算することで、容易に高い並列度で並列処理することができ、検索速度をさらに向上させることが可能である。

産業上の利用可能性

以上のように本発明のベクトル索引作成方法は、部分ベクトル算定手段と、ノルム分布集計手段と、ノルム区分表と、領域番号算定手段と、
10 偏角分布集計手段と、偏角区分表と、ノルム区分番号算定手段と、偏角区分番号算定手段と、索引データ算定手段と、索引構成手段とを有することにより、方向、ノルムの分布が不明なベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、類似検索の際に、ベクトル間の距離と、ベクトルの内積の2種類の類似度の
15 いずれかを選択でき、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「内積が0.6以上」といった、類似検索範囲の指定が可能で、索引作成に要する計算量が実用的範囲であるような、ベクトル索引の作成を可能にする、という効果を奏する
20 ものである。

また、本発明のベクトル索引作成方法は、成分区分番号算定手段をさらに有することにより、上記の効果に加えて、成分の量子化による計算誤差を最小限に抑えながら、作成すべきベクトル索引の容量を大幅に低減できる、という効果を奏するものである。

25 また、本発明の類似ベクトル検索方法は、部分質問条件算定手段と、検索対象範囲生成手段と、索引検索手段と、内積差分上限算定手段もし

くは二乗距離差分上限算定手段と、類似検索結果決定手段とを有し、部分内積差分の累算値を算定して類似検索の手がかりとして用いることにより、ベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、「もっとも類似している L 個のベクトルを求める」という型の類似検索ができ、さらに、 L が比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「内積が 0.6 以上」といった、類似検索範囲の指定が可能な、内積あるいは距離を類似尺度として用いる類似ベクトル検索を可能にする、という効果を奏するものである。しかも、内積と距離のどちらを類似尺度として用いるかをベクトル索引の作成時に指定する必要がなく、単一のベクトル索引を使って、検索時に必要に応じて、類似尺度の使い分けが可能であるという、優れた効果を奏するものである。

また、本発明の類似ベクトル検索方法は、部分質問条件を算定する手段と、検索対象範囲を生成する手段と、索引を検索する手段と、二乗距離差分上限を算定する手段と、類似検索結果を決定する手段とを有し、部分二乗距離差分の累算値を算定して類似検索の手がかりとして用いることにより、ベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、「もっとも類似している L 個のベクトルを求める」という型の類似検索ができ、さらに、 L が比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「距離が 0.8 以下」といった、類似検索範囲の指定が可能で、距離を類似尺度として用いる類似ベクトル検索を可能にする、という効果を奏するものである。

本発明の効果は、索引作成対象、検索対象となるベクトルデータの次元が数百次元と高く、ベクトルデータベース中のベクトルデータ数が数十万～数百万個と多く、検索時の取得件数が数十個と多い場合には特に

顕著であり、従来のベクトル索引作成方法では数百時間を要していた索引作成時間を数十分に短縮するとともに、従来の類似ベクトル検索方法では数分を要する、あるいは実行不可能であった類似検索処理を1秒以下で処理することができ、実用上、非常に大きな効果が得られる。

請 求 の 範 囲

1. 少なくともN次元実ベクトルと、その識別番号とを含む組をベクトルデータとして有限個登録したベクトルデータベースに対して、
- 5 機械検索可能な索引を作成する方法であって、

- ベクトル索引作成の第1ステップでは、前記ベクトルデータベース中の各ベクトルデータのN次元実ベクトルVに対して、N個の成分を予め定めた方法でm組に分けて、m個の部分ベクトル $v_1 \sim v_m$ を作成した後、部分ベクトル v_k ($k = 1 \sim m$) のノルムの分布を集計して、予め定め
- 10 られたD種のノルム区分のノルムの範囲を定めたノルム区分表を作成するとともに、予め定めたD個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_k の所属する領域番号dを算定し、前記部分ベクトル v_k と領域中心ベクトル p_d とのなす角の余弦 $(v_k \cdot p_d) / (|v_k| * |p_d|)$ の分布を偏角分布として集計して、予め定めたC種の偏角
- 15 区分の偏角の範囲を記録した偏角区分表を作成し、

- ベクトル索引作成の第2ステップでは、前記ベクトルデータベース中の各ベクトルデータのN次元実ベクトルVに対して、前記第1のステップと同一の方法でN個の成分をm組に分けて、m個の部分ベクトル $v_1 \sim v_m$ を作成した後、部分空間番号bに対する部分ベクトル v_b ($b = 1$
- 20 $\sim m$) に対して、前記ノルム区分表を参照して前記各部分ベクトル v_b のノルムが所属するノルム区分の番号rを算定するとともに、前記第1のステップと同一の方法で、予め定めたD個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_b の所属する領域番号dを算定し、前記部分ベクトル v_b と前記領域番号dの領域の中心方向を表す領域中心
- 25 ベクトル p_d とのなす角の余弦である偏角 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ を算定して、前記偏角区分表を参照して、所属する偏角区分の

番号 c を算出し、前記部分空間番号 b と、前記領域番号 d と、前記偏角区分番号 c と、前記ノルム区分番号 r と、前記部分ベクトル v_b の成分と、識別番号 i とから、ベクトル索引に登録する索引登録データを算定し、

- 5 ベクトル索引作成の第3ステップでは、前記ノルム区分表と、前記偏角区分表と、前記索引登録データとから、各部分ベクトルの識別番号と成分とを、部分空間番号 b 、領域番号 d 、角区分番号 c 、ノルム区分番号範囲 $[r_1, r_2]$ の組を鍵にして、索可能で、かつ各ベクトルデータのベクトルの成分を、その識別番号で検索可能なベクトル索引として構成
- 10 成することを特徴とするベクトル索引作成方法。

2. 少なくとも N 次元実ベクトルと、その識別番号とを含む組をベクトルデータとして有限個登録したベクトルデータベースに対して、機械検索可能な索引を作成する方法であって、

- 15 ベクトル索引作成の第1ステップでは、前記ベクトルデータベース中の各ベクトルデータの N 次元実ベクトル V に対して、 N 個の成分を予め定めた方法で m 組に分けて、 m 個の部分ベクトル $v_1 \sim v_m$ を作成した後、各部分空間番号 b に対する部分ベクトル v_b ($b = 1 \sim m$) のノルムの分布を集計して、予め定められた D 種のノルム区分のノルムの範囲を定
- 20 めたノルム区分表を作成するとともに、予め定めた D 個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_b の所属する領域番号 d を算定し、前記部分ベクトル v_b と領域中心ベクトル p_d とのなす角の余弦 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ の分布を偏角分布として集計して、予め定めた C 種の偏角区分の偏角の範囲を記録した偏角区分表を作成し、
- 25 ベクトル索引作成の第2ステップでは、前記ベクトルデータベース中の各ベクトルデータの N 次元実ベクトル V に対して、前記第1ステップ

- と同一の方法でN個の成分をm組に分けて、m個の部分ベクトル $v_1 \sim v_m$ を作成した後、前記部分空間 b に対する部分ベクトル v_b ($b = 1 \sim m$) に対して、前記ノルム区分表を参照して前記各部分ベクトル v_b のノルムが所属するノルム区分の番号 r を算定するとともに、前記第1ステップと同一の方法で、予め定めたD個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_b の所属する領域番号 d を算定し、前記部分ベクトル v_b と前記領域番号 d の領域の中心方向を表す領域中心ベクトル p_d とのなす角の余弦である偏角 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ を算定して、前記偏角区分表を参照して、所属する偏角区分の番号 c を算出し、前記算定したノルム区分番号 r に対応するノルム区分のノルムの最高値から、前記算定した部分ベクトル v_b の各成分 v_{bj} に対して、 v_{bj} が所属する予め定めた範囲の成分区分番号 w_j を算定し、前記部分空間番号 b と、前記領域番号 d と、前記偏角区分番号 c と、前記ノルム区分番号 r と、前記成分区分番号 w_j の列と、識別番号 i とから、
- ベクトル索引に登録する索引登録データを算定し、

- ベクトル索引作成の第3ステップでは、前記ノルム区分表と、前記偏角区分表と、前記索引登録データとから、各部分ベクトルの識別番号と成分とを、部分空間番号 b 、領域番号 d 、偏角区分番号 c 、ノルム区分番号範囲 $[r_1, r_2]$ の組を鍵にして、検索可能で、かつ各ベクトルデータのベクトルの成分を、その識別番号で検索可能なベクトル索引として構成することを特徴とするベクトル索引作成方法。

3. 前記ベクトル索引作成の第1および第2のステップにおいて、部分ベクトル v_b と領域中心ベクトル p_d とのなす角度の関数として、角度の余弦 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ を用い、この値を偏角として偏角分布を求める、請求項1又は2記載のベクトル索引作成方法。

4. 前記ベクトル索引作成の第1および第2のステップにおいて、
N次元ベクトルVの全成分を取り出すように、Vの先頭の成分から順に、
N/m個、あるいは $(N/m) + 1$ 個の成分を取り出して、部分ベクトル
5 を作成する請求項1又は2記載のベクトル索引作成方法。

5. 前記ベクトル索引作成の第1のステップにおいて、ノルム区分表を作成する際に、ノルム分布の集計結果をもとに、各ノルム区分に相当するノルム範囲に所属する部分ベクトルの数が、できる限り均一になるようにノルム区分を定める請求項1記載のベクトル索引作成方法。
10

6. 前記ベクトル索引作成の第1のステップにおいて、偏角区分表を作成する際に、偏角分布の集計結果をもとに、各偏角区分に相当する偏角範囲に所属する部分ベクトルの数が、できる限り均一になるように偏角区分を定める請求項1記載のベクトル索引作成方法。
15

7. 前記ベクトル索引作成の第1および第2のステップにおいて、部分ベクトル v_b の領域番号を、予め定めたD個の領域中心ベクトル $p_1 \sim p_D$ のうち、 p_d と v_b のなす角の余弦 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ がもっとも大きくなるような領域中心ベクトル p_d の番号dとして求める請求項1又は2記載のベクトル索引作成方法。
20

8. 前記ベクトル索引作成の第3のステップにおいて、部分空間番号b、領域番号d、偏角区分番号c、ノルム区分番号rを結合した番号 $(b * Nd * Nc * Nr) + (d * Nc * Nr) + (c * Nr) + r$ を鍵とし、ベクトルの識別番号iと成分とを検索可能な探索木と、ベクトルデ
25

ータの識別番号を添字とする、各部分ベクトルの前記探索木の鍵を記録した表とを作成し、ベクトル索引の一部とする請求項 1 又は 2 記載のベクトル索引作成方法。

- 5 9. 前記ベクトル索引作成の第 2 のステップにおいて、成分が $\{-1, 0, +1\}$ のうちのいずれかである、0 ベクトルでない全てのベクトル $(0, \dots, 0, +1) \sim (-1, \dots, -1)$ を正規化したベクトルを領域中心ベクトルとして用いる請求項 1 又は 2 記載のベクトル索引作成方法。

10

10. 検索条件として N 次元実ベクトルの質問ベクトル Q と、内積下限値 α 、および最大取得ベクトル数 L とを指定して、少なくとも N 次元実ベクトルとその ID 番号との組を有限個登録したベクトルデータから作成したベクトル索引を検索して、前記ベクトルデータベースのベクトルデータ (i, V) で、前記質問ベクトル Q との内積の値 $V \cdot Q$ が前記内積下限値 α より大きいものについて、その識別番号 i と、 Q と V の内積の組 $(i, V \cdot Q)$ を最大 L 個求める類似ベクトルの検索方法であって、

- 類似ベクトル検索の第 1 ステップでは、前記質問ベクトル Q に対して、
20 Q の N 個の成分を、前記ベクトル索引の作成時に用いた方法と同一の、予め定めた方法で m 組に分けて、 m 個の部分質問ベクトル $q_1 \sim q_m$ を作成するとともに、各部分質問ベクトル q_b と対応する部分ベクトルとの内積（この内積をこれ以降「部分内積」と呼ぶ）の下限値である部分内積下限値 f_b を、指定された内積下限値 α から算定し、各部分質問ベクトル q_b ($b = 1 \sim m$) と各領域 b に対して、領域中心ベクトル p_d と前記部分質問ベクトル q_b との内積 $p_d \cdot q_b$ の値と、前記部分内積下限値

- f_b と、前記ベクトル索引中のノルム区分表および偏角区分表とから、部分空間番号 b 、領域番号 d において検索すべき偏角区分番号 c 、ノルム区分の範囲 $[r_1, r_2]$ の組 $(c, [r_1, r_2])$ を算定し、前記算定した $(c, [r_1, r_2])$ を基に、 $(b, d, c, [r_1, r_2])$ を
- 5 検索条件として前記ベクトル索引を範囲検索し、索引検索結果として、条件を満たす部分ベクトル v_b の識別番号 i と成分とを求め、前記 v_b と q_b との部分内積 $v_b \cdot q_b$ と前記部分内積下限値 f_b との差である、部分内積差分 $(v_b \cdot q_b) - f_b$ を算出して、内積差分表の識別番号 i の内積差分上限値 $S[i]$ として累算（加算）し、
- 10 類似ベクトル検索の第2ステップでは、前記内積差分表 $S[i]$ 中で値の大きいものから順に、前記ベクトル索引を識別番号 i で検索してベクトルデータの成分 V を求め、 V と前記質問ベクトル Q との内積 $V \cdot Q$ から α を差し引いた内積差分値 $t = V \cdot Q - \alpha$ を算出していき、前記内積差分表中で、内積差分値を算出していない要素の最大値より大きい内積
- 15 差分値を持つベクトルデータが L 個以上になった時点、あるいは内積差分上限値が正である全てのベクトルデータの内積を算出した時点で、内積差分値の大きな最大 L 個のベクトルデータに対し、少なくとも識別番号 i と内積 $t + \alpha$ との組を、検索結果として出力することを特徴とする類似ベクトル検索方法。
- 20
- 1 1. 検索条件として N 次元実ベクトルの質問ベクトル Q と、距離上限値 α 、および最大取得ベクトル数 L とを指定して、少なくとも N 次元実ベクトルとその識別番号との組を有限個登録したベクトルデータから作成したベクトル索引を検索して、前記質問ベクトル Q との内積の値
- 25 が前記距離上限値 α 以下であるような、前記ベクトルデータ中の N 次元実ベクトル V の識別番号 i と、 Q と V の距離 p の組 (i, p) を最大 L

個求める類似ベクトルの検索方法であって、

- 類似ベクトル検索の第1ステップでは、前記質問ベクトルQに対して、QのN個の成分を、前記ベクトル索引の作成時に用いた方法と同一の、予め定めた方法でm組に分けて、m個の部分質問ベクトル $q_1 \sim q_m$ を作成するとともに、各部分質問ベクトル q_b と対応する部分ベクトル v_b との二乗距離 $|v_b - q_b|^2$ （すなわちユークリッド距離の二乗、この値をこれ以降「部分二乗距離」と呼ぶ）の上限値である部分二乗距離上限値 f_b を、指定された距離上限値 α から算定し、各部分質問ベクトル q_b （ $b = 1 \sim m$ ）に対して、前記部分質問ベクトル q_b と、前記部分二乗距離上限値 f_b と、前記ベクトル索引中のノルム区分表および偏角区分表とから、検索すべき部分空間番号b、領域番号d、偏角区分番号c、ノルム区分の範囲 $[r_1, r_2]$ の組（ $b, d, c, [r_1, r_2]$ ）を系統的に生成し、前記生成した（ $b, d, c, [r_1, r_2]$ ）を検索条件として、前記ベクトル索引を範囲検索して、索引検索結果として、条件を満たす部分ベクトル v_b の識別番号iと成分とを求め、前記部分二乗距離上限値 f_b と、 v_b と q_b との部分二乗距離 $|v_b - q_b|^2$ との差である、部分二乗距離差分 $f_b - |v_b - q_b|^2$ を算出して、二乗距離差分表の識別番号iの二乗距離差分上限値 $S[i]$ として累算（加算）し、

- 類似ベクトル検索の第2ステップでは、前記二乗距離差分表 $S[i]$ 中で値の大きいものから順に、前記ベクトル索引を識別番号iで検索してベクトルデータの成分Vを求め、距離上限値の二乗 α^2 からVと前記質問ベクトルQとの二乗距離 $|V - Q|^2$ を差し引いた二乗距離差分値 $\alpha^2 - |V - Q|^2$ を算出していき、前記二乗距離差分表中で、二乗距離差分値を算出していない要素の最大値より大きい二乗距離差分値を持つベクトルデータがL個以上になった時点、あるいは二乗距離差分上限値が正である全てのベクトルデータの二乗距離差分値を算出した時点で、二

乗距離差分値 t の大きな最大 L 個のベクトルデータに対し、少なくとも識別番号 i と、距離 $(\alpha^2 - t)^{1/2}$ との組を、検索結果として出力することを特徴とする類似ベクトル検索方法。

- 5 1 2. 前記類似ベクトル検索の第 1 のステップにおいて、 N 次元ベクトル V の全成分を取り出すように、 V の先頭の成分から順に、 N/m 個、あるいは $(N/m) + 1$ 個の成分を取り出して、部分質問ベクトルを作成する請求項 1 0 又は 1 1 記載の類似ベクトル検索方法。
- 10 1 3. 前記類似ベクトル検索の第 1 のステップにおいて、前記部分質問ベクトル q_b と対応する部分ベクトル v_b との内積の下限値である部分内積下限値 f_b を、指定された内積下限値 α から、 $f_b = \alpha |q_b|^2 / \Sigma(|q_b|^2)$ によって算定する請求項 1 1 記載の類似ベクトル検索方法。
- 15
- 1 4. 前記類似ベクトル検索の第 1 のステップにおいて、前記部分質問ベクトル q_b と対応する部分ベクトル v_b との二乗距離の上限値である部分二乗距離上限値 f_b を、指定された距離下上限値 α から、 $f_b = \alpha^2 |q_b|^2 / \Sigma(|q_b|^2)$ によって算定する請求項 1 1 記載の類似
- 20 ベクトル検索方法。
- 1 5. 少なくとも N 次元実ベクトルと、その識別番号とを含む組をベクトルデータとして有限個登録したベクトルデータベースに対して、機械検索可能な索引を作成する装置であって、
- 25 前記ベクトルデータベース中の各ベクトルデータの N 次元実ベクトル V に対して、 N 個の成分を予め定めた方法で m 組に分けて、 m 個の部分

ベクトル $v_1 \sim v_m$ を作成する部分ベクトル算定手段と、

- 前記作成された m 個の部分ベクトル $v_1 \sim v_m$ のうち、部分ベクトル v_k ($k = 1 \sim m$) のノルムの分布を集計して、予め定められた D 種のノルム区分のノルムの範囲を定めたノルム区分表を作成するノルム分布集計手段と、

予め定めた D 個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_k の所属する領域番号 d を算定する領域番号算定手段と、

- 前記部分ベクトル v_k と領域中心ベクトル p_d とのなす角の余弦 ($v_k \cdot p_d$) / ($|v_k| * |p_d|$) の分布を偏角分布として集計して、予め定めた C 種の偏角区分の偏角の範囲を記録した偏角区分表を作成する偏角分布集計手段と、

- 前記部分ベクトル算定手段が作成した m 個の部分ベクトル $v_1 \sim v_m$ のうち、部分空間番号 b に対する部分ベクトル v_b ($b = 1 \sim m$) に対して、前記ノルム区分表を参照して前記各部分ベクトル v_b のノルムが所属するノルム区分の番号 r を算定するノルム区分番号算定手段と、

- 前記部分ベクトル v_b と前記領域番号算定手段が算定した領域番号 d の領域の中心方向を表す領域中心ベクトル p_d とのなす角の余弦である偏角 ($v_b \cdot p_d$) / ($|v_b| * |p_d|$) を算定して、前記偏角区分表を参照して、所属する偏角区分の番号 c を算出する偏角区分番号算定手段と、

- 前記部分空間番号 b と、前記領域番号 d と、前記偏角区分番号 c と、前記ノルム区分番号 r と、前記部分ベクトル v_b の成分と、識別番号 i とから、ベクトル索引に登録する索引登録データを算定する索引データ算定手段と、

- 前記ノルム区分表と、前記偏角区分表と、前記索引登録データとから、各部分ベクトルの識別番号と成分とを、部分空間番号 b 、領域番号 d 、

偏角区分番号 c 、ノルム区分番号範囲 $[r_1, r_2]$ の組を鍵にして、検索可能で、かつ各ベクトルデータのベクトルの成分を、その識別番号で検索可能なベクトル索引として構成する索引構成手段とを備えたことを特徴とするベクトル索引作成装置。

5

16. 少なくとも N 次元実ベクトルと、その識別番号とを含む組をベクトルデータとして有限個登録したベクトルデータベースに対して、機械検索可能な索引を作成する装置であって、

前記ベクトルデータベース中の各ベクトルデータの N 次元実ベクトル V に対して、 N 個の成分を予め定めた方法で m 組に分けて、 m 個の部分ベクトル $v_1 \sim v_m$ を作成する部分ベクトル算定手段と、

前記作成された m 個の部分ベクトル $v_1 \sim v_m$ のうち、部分空間番号 b に対する部分ベクトル v_b ($b = 1 \sim m$) のノルムの分布を集計して、予め定められた D 種のノルム区分のノルムの範囲を定めたノルム区分表

15 作成するノルム分布集計手段と、

予め定めた D 個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_b の所属する領域番号 d を算定する領域番号算定手段と、

前記部分ベクトル v_b と領域中心ベクトル p_d とのなす角の余弦 ($v_b \cdot p_d$) / ($|v_b| * |p_d|$) の分布を偏角分布として集計して、予め定めた C 種の偏角区分の偏角の範囲を記録した偏角区分表を作成する

20 偏角分布集計手段と、

前記部分ベクトル算定手段が作成した m 個の部分ベクトル $v_1 \sim v_m$ のうち、部分空間 b に対する部分ベクトル v_b ($b = 1 \sim m$) に対して、前記ノルム区分表を参照して前記各部分ベクトル v_b のノルムが所属するノルム区分の番号 r を算定するノルム区分番号算定手段と、

25

前記部分ベクトル v_b と前記領域番号算定手段が算定した領域番号 d

の領域の中心方向を表す領域中心ベクトル p_d とのなす角の余弦である偏角 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ を算定して、前記偏角区分表を参照して、所属する偏角区分の番号 c を算出する偏角区分番号算定手段と、

- 5 前記算定したノルム区分番号 r に対応するノルム区分のノルムの最高値から、前記算定した部分ベクトル v_b の各成分 v_{b_j} に対して、 v_{b_j} が所属する予め定めた範囲の成分区分番号 w_j を算定する成分区分番号算定手段と、

- 前記部分空間番号 b と、前記領域番号 d と、前記偏角区分番号 c と、
10 前記ノルム区分番号 r と、前記成分区分番号 w_j の列と、識別番号 i とから、ベクトル索引に登録する索引登録データを算定する索引データ算定手段と、

- 前記ノルム区分表と、前記偏角区分表と、前記索引登録データとから、各部分ベクトルの識別番号と成分とを、部分空間番号 b 、領域番号 d 、
15 偏角区分番号 c 、ノルム区分番号範囲 $[r_1, r_2]$ の組を鍵にして、検索可能で、かつ各ベクトルデータのベクトルの成分を、その識別番号で検索可能なベクトル索引として構成する索引構成手段とを備えたことを特徴とするベクトル索引作成装置。

- 20 17. 前記部分ベクトル算定手段が、 N 次元ベクトル V の全成分を取り出すように、 V の先頭の成分から順に、 N/m 個、あるいは $(N/m) + 1$ 個の成分を取り出して、部分ベクトルを作成する請求項 15 又は 16 記載のベクトル索引作成装置。

- 25 18. 前記ノルム分布集計手段が、ノルム区分表を作成する際に、ノルム分布の集計結果をもとに、各ノルム区分に相当するノルム範囲に

所属する部分ベクトルの数が、できる限り均一になるようにノルム区分を定める請求項 15 記載のベクトル索引作成装置。

19. 前記偏角分布集計手段が、偏角区分表を作成する際に、偏角
5 分布の集計結果をもとに、各偏角区分に相当する偏角範囲に所属する部分ベクトルの数が、できる限り均一になるように偏角区分を定める請求項 15 記載のベクトル索引作成装置。
20. 前記領域番号算定手段が、部分ベクトル v_b の領域番号を、
10 予め定めた D 個の領域中心ベクトル $p_1 \sim p_D$ のうち、 p_d と v_b のなす角の余弦 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ がもっとも大きくなるような領域中心ベクトル p_d の番号 d として求める請求項 15 又は 16 記載のベクトル索引作成装置。
- 15 21. 前記索引構成手段が、部分空間番号 b 、領域番号 d 、偏角区分番号 c 、ノルム区分番号 r を結合した番号 $(b * Nd * Nc * Nr) + (d * Nc * Nr) + (c * Nr) + r$ を鍵とし、ベクトルの識別番号 i と成分とを検索可能な探索木と、ベクトルデータの識別番号を添字とする、各部分ベクトルの前記探索木の鍵を記録した表とを作成し、ベクトル索引
20 引の一部とする請求項 15 又は 16 記載のベクトル索引作成装置。
22. 前記領域番号算定手段が、成分が $\{-1, 0, +1\}$ のうちのいずれかである、0 ベクトルでない全てのベクトル $(0, \dots, 0, +1) \sim (-1, \dots, -1)$ を正規化したベクトルを領域中心ベクトル
25 として用いる請求項 15 又は 16 記載のベクトル索引作成装置。

23. 検索条件としてN次元実ベクトルの質問ベクトルQと、内積下限値 α 、および最大取得ベクトル数Lとを指定して、少なくともN次元実ベクトルとそのID番号との組を有限個登録したベクトルデータから作成したベクトル索引を検索して、前記ベクトルデータベースのベクトルデータ(i, V)で、前記質問ベクトルQとの内積の値 $V \cdot Q$ が前記内積下限値 α より大きいものについて、その識別番号iと、QとVの内積の組(i, $V \cdot Q$)を最大L個求める類似ベクトルの検索装置であって、

前記質問ベクトルQに対して、QのN個の成分を、前記ベクトル索引の作成時に用いた方法と同一の、予め定めた方法でm組に分けて、m個の部分質問ベクトル $q_1 \sim q_m$ を作成するとともに、各部分質問ベクトル q_b と対応する部分ベクトルとの内積(この内積をこれ以降「部分内積」と呼ぶ)の下限値である部分内積下限値 f_b を、指定された内積下限値 α から算定する部分質問条件算定手段と、

15 前記部分質問ベクトル q_b ($b = 1 \sim m$) と各領域bに対して、領域中心ベクトル p_d と前記部分質問ベクトル q_b との内積 $p_d \cdot q_b$ の値と、前記部分内積下限値 f_b と、前記ベクトル索引中のノルム区分表および偏角区分表とから、部分空間番号b、領域番号dにおいて検索すべき偏角区分番号c、ノルム区分の範囲 $[r_1, r_2]$ の組(c, $[r_1, r_2]$)を算定する検索対象範囲生成手段と、

前記検索対象範囲生成手段が算定した(c, $[r_1, r_2]$)を基に、(b, d, c, $[r_1, r_2]$)を検索条件として前記ベクトル索引を範囲検索し、索引検索結果として、条件を満たす部分ベクトル v_b の識別番号iと成分とを求める索引検索手段と、

25 前記 v_b と q_b との部分内積 $v_b \cdot q_b$ と前記部分内積下限値 f_b との差である、部分内積差分($v_b \cdot q_b$) - f_b を算出して、内積差分表の識

別番号 i の内積差分上限値 $S[i]$ として累算（加算）する内積差分上限算定手段と、

- 前記内積差分表 $S[i]$ 中で値の大きいものから順に、前記ベクトル索引を識別番号 i で検索してベクトルデータの成分 V を求め、 V と前記質問ベクトル Q との内積 $V \cdot Q$ から α を差し引いた内積差分値 $t = V \cdot Q - \alpha$ を算出していき、前記内積差分表中で、内積差分値を算出していない要素の最大値より大きい内積差分値を持つベクトルデータが L 個以上になった時点、あるいは内積差分上限値が正である全てのベクトルデータの内積を算出した時点で、内積差分値の大きな最大 L 個のベクトルデータに対し、少なくとも識別番号 i と内積 $t + \alpha$ との組を、検索結果として出力する類似検索結果決定手段とを備えたことを特徴とする類似ベクトル検索装置。

24. 検索条件として N 次元実ベクトルの質問ベクトル Q と、距離上限値 α 、および最大取得ベクトル数 L とを指定して、少なくとも N 次元実ベクトルとその識別番号との組を有限個登録したベクトルデータから作成したベクトル索引を検索して、前記質問ベクトル Q との内積の値が前記距離上限値 α 以下であるような、前記ベクトルデータ中の N 次元実ベクトル V の識別番号 i と、 Q と V の距離 p の組 (i, p) を最大 L 個求める類似ベクトルの検索装置であって、

- 前記質問ベクトル Q に対して、 Q の N 個の成分を、前記ベクトル索引の作成時に用いた方法と同一の、予め定めた方法で m 組に分けて、 m 個の部分質問ベクトル $q_1 \sim q_m$ を作成するとともに、各部分質問ベクトル q_b と対応する部分ベクトル v_b との二乗距離 $|v_b - q_b|^2$ （すなわちユークリッド距離の二乗、この値のことをこれ以降「部分二乗距離」と呼ぶ）の上限値である部分二乗距離上限値 f_b を、指定された距離上限

値 α から算定する部分質問条件算定手段と、

前記部分質問ベクトル q_b ($b = 1 \sim m$) に対して、前記部分質問ベクトル q_b と、前記部分二乗距離上限値 f_b と、前記ベクトル索引中のノルム区分表および偏角区分表とから、検索すべき部分空間番号 b 、領域

5 番号 d 、偏角区分番号 c 、

ノルム区分の範囲 $[r_1, r_2]$ の組 $(b, d, c, [r_1, r_2])$ を系統的に生成する検索対象範囲生成手段と、

前記検索対象範囲生成手段が生成した $(b, d, c, [r_1, r_2])$ を検索条件として、前記ベクトル索引を範囲検索して、索引検索結果として、条件を満たす部分ベクトル v_b の識別番号 i と成分とを求める索引検索手段と、

10 前記部分二乗距離上限値 f_b と、 v_b と q_b との部分二乗距離 $|v_b - q_b|^2$ との差である、部分二乗距離差分 $f_b - |v_b - q_b|^2$ を算出して、二乗距離差分表の識別番号 i の二乗距離差分上限値 $S[i]$ として累算

15 (加算) する二乗距離差分上限算定手段と、

前記二乗距離差分表 $S[i]$ 中で値の大きいものから順に、前記ベクトル索引を識別番号 i で検索してベクトルデータの成分 V を求め、距離上限値の二乗 α^2 から V と前記質問ベクトル Q との二乗距離 $|V - Q|^2$ を差し引いた二乗距離差分値 $\alpha^2 - |V - Q|^2$ を算出していき、前記二乗距離差分表中で、二乗距離差分値を算出していない要素の最大値より大きい二乗距離差分値を持つベクトルデータが L 個以上になった時点、あるいは二乗距離差分上限値が正である全てのベクトルデータの二乗距離差分値を算出した時点で、二乗距離差分値 t の大きな最大 L 個のベクトルデータに対し、少なくとも識別番号 i と距離 $(\alpha^2 - t)^{1/2}$ との組

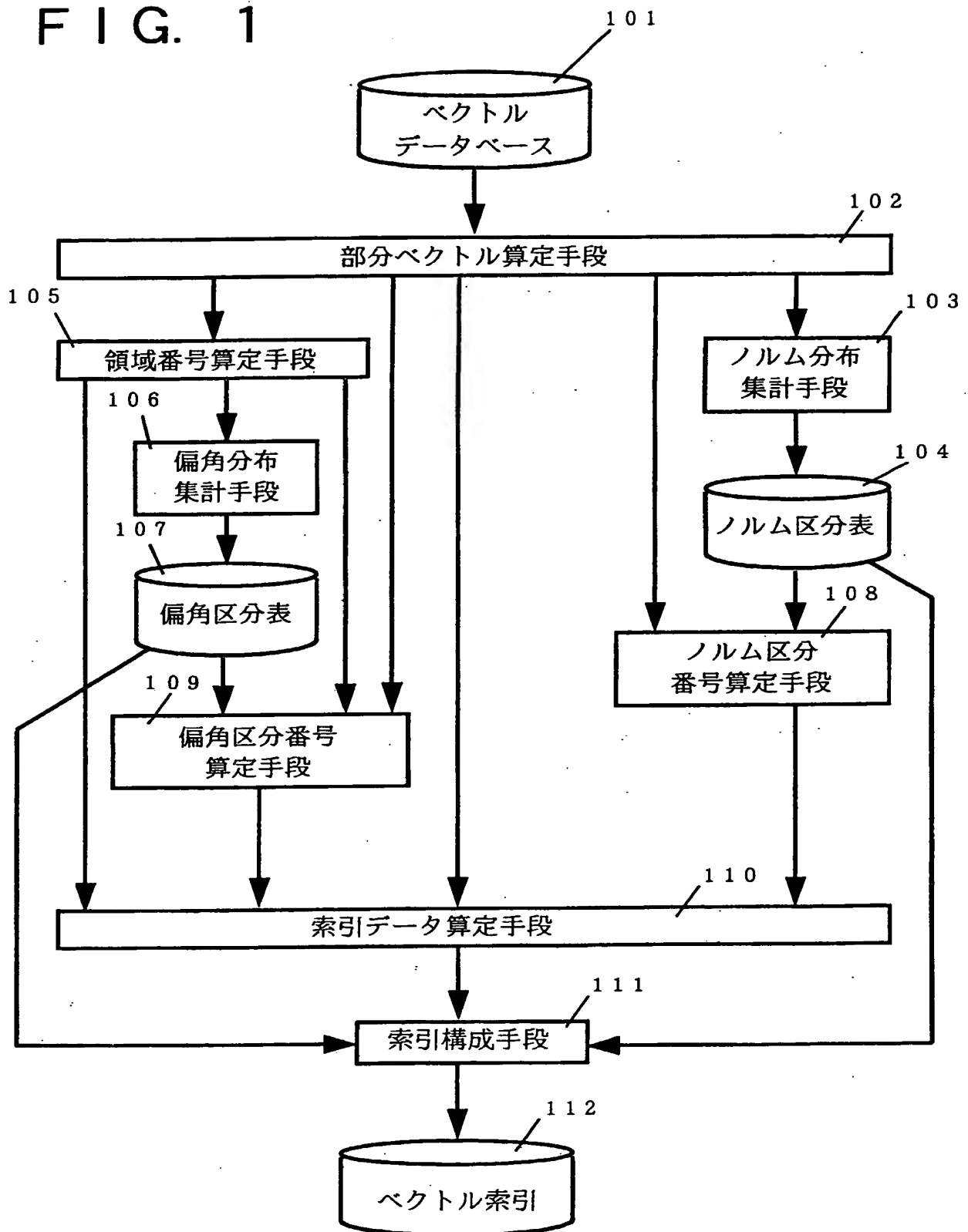
20

25 を、検索結果として出力する類似検索結果決定手段とを備えたことを特徴とする類似ベクトル検索装置。

25. 前記部分質問条件算定手段が、N次元ベクトルVの全成分を取り出すように、Vの先頭の成分から順に、 N/m 個、あるいは $(N/m) + 1$ 個の成分を取り出して、部分質問ベクトルを作成する請求項25 3又は24記載の類似ベクトル検索装置。
26. 前記部分質問ベクトル q_b と対応する部分ベクトル v_b との内積の下限値である部分内積下限値 f_b を、指定された内積下限値 α から、 $f_b = \alpha |q_b|^2 / \sum (|q_b|^2)$ によって算定する請求項23記載の10 類似ベクトル検索装置。
27. 前記部分質問ベクトル q_b と対応する部分ベクトル v_b との二乗距離の上限値である部分二乗距離上限値 f_b を、指定された距離下上限値 α から、 $f_b = \alpha^2 |q_b|^2 / \sum (|q_b|^2)$ によって算定する請求項24記載の類似ベクトル検索装置。15
28. 請求項1又は2記載の方法を実行するためのコンピュータプログラムを記録した記録媒体。
29. 請求項15又は16記載の装置をソフトウェアにより実現するためのコンピュータプログラムを記録した記録媒体。20

1/29

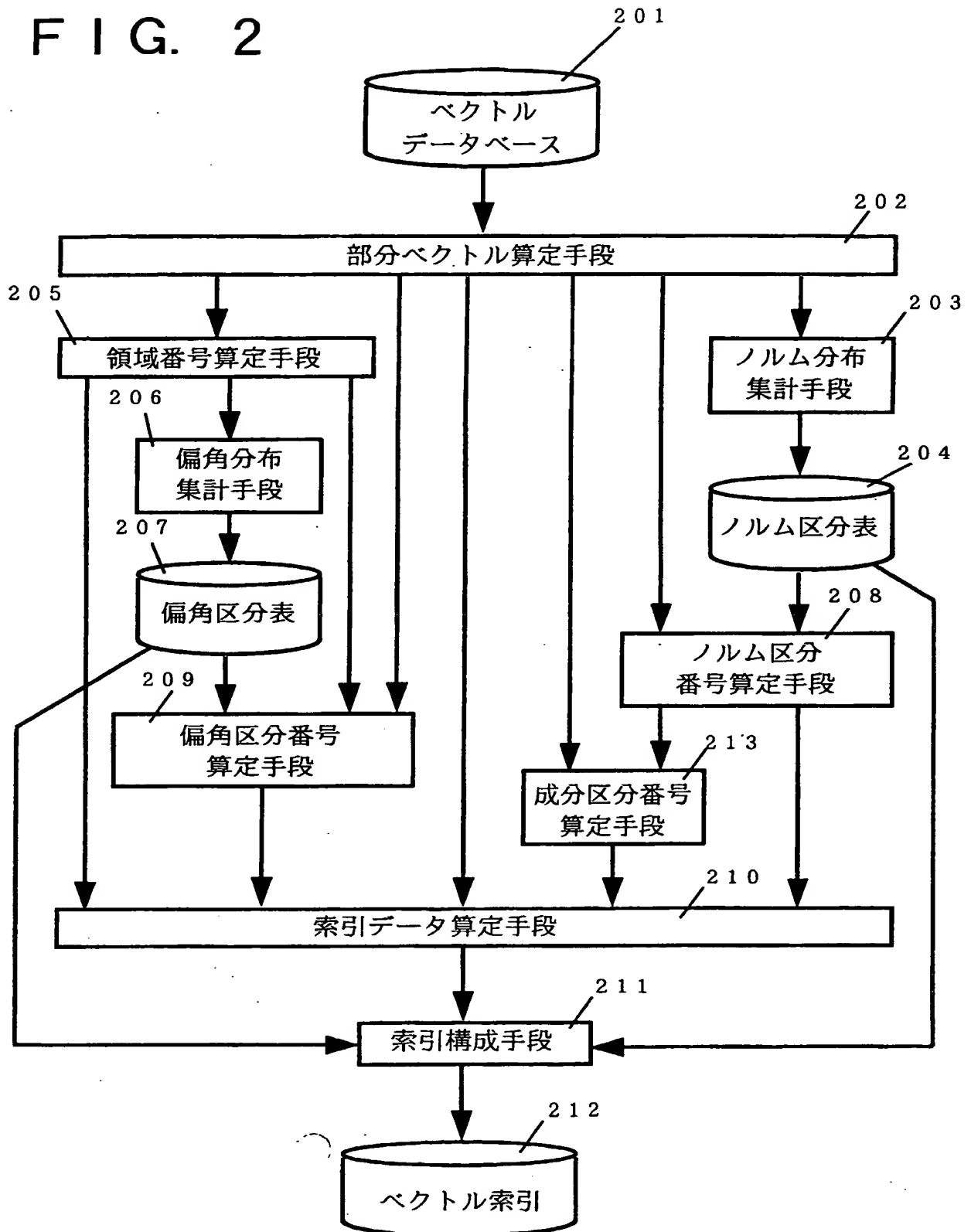
FIG. 1



THIS PAGE BLANK (USPTO)

2/29

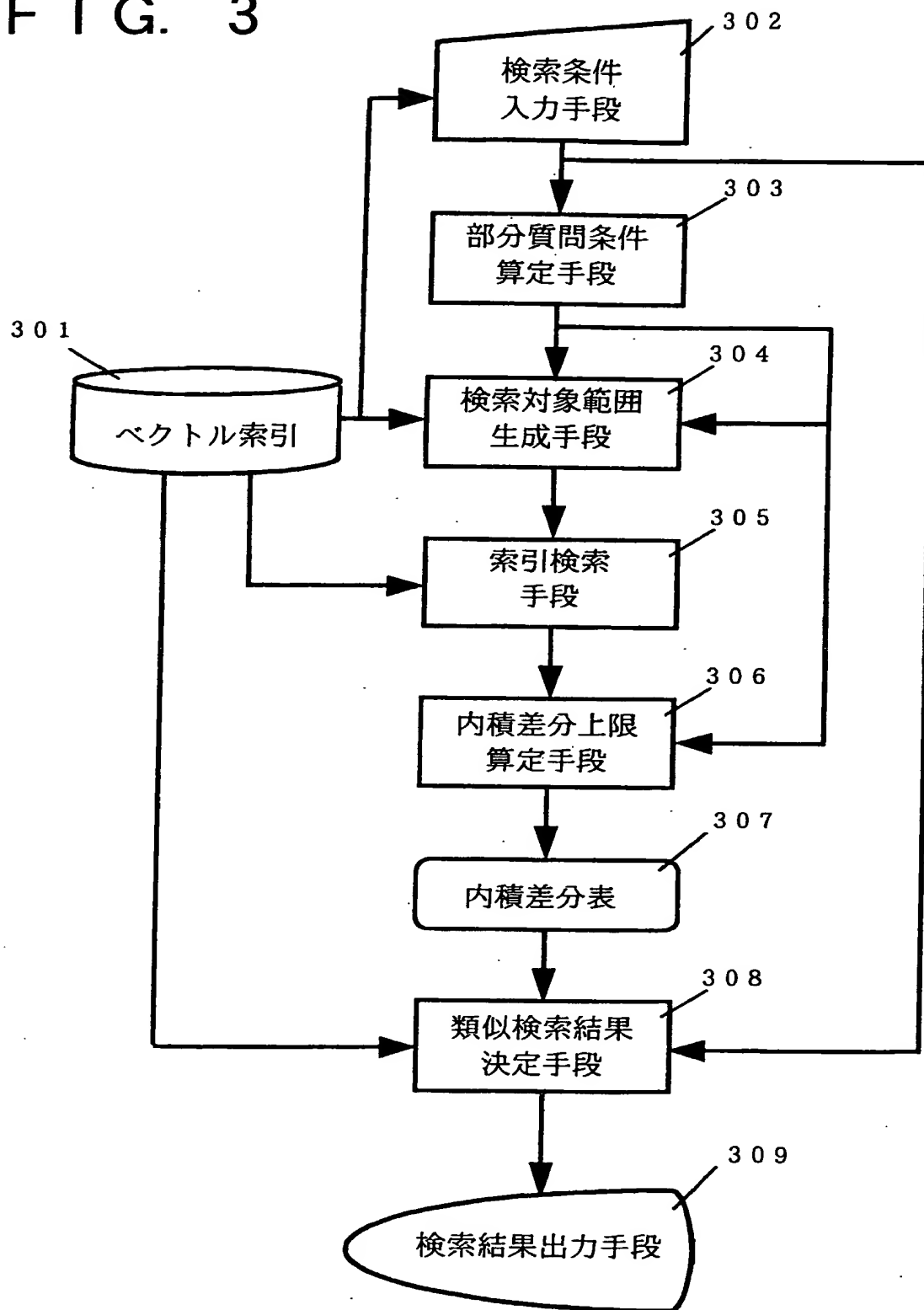
FIG. 2



THIS PAGE BLANK (USPTO)

3/29

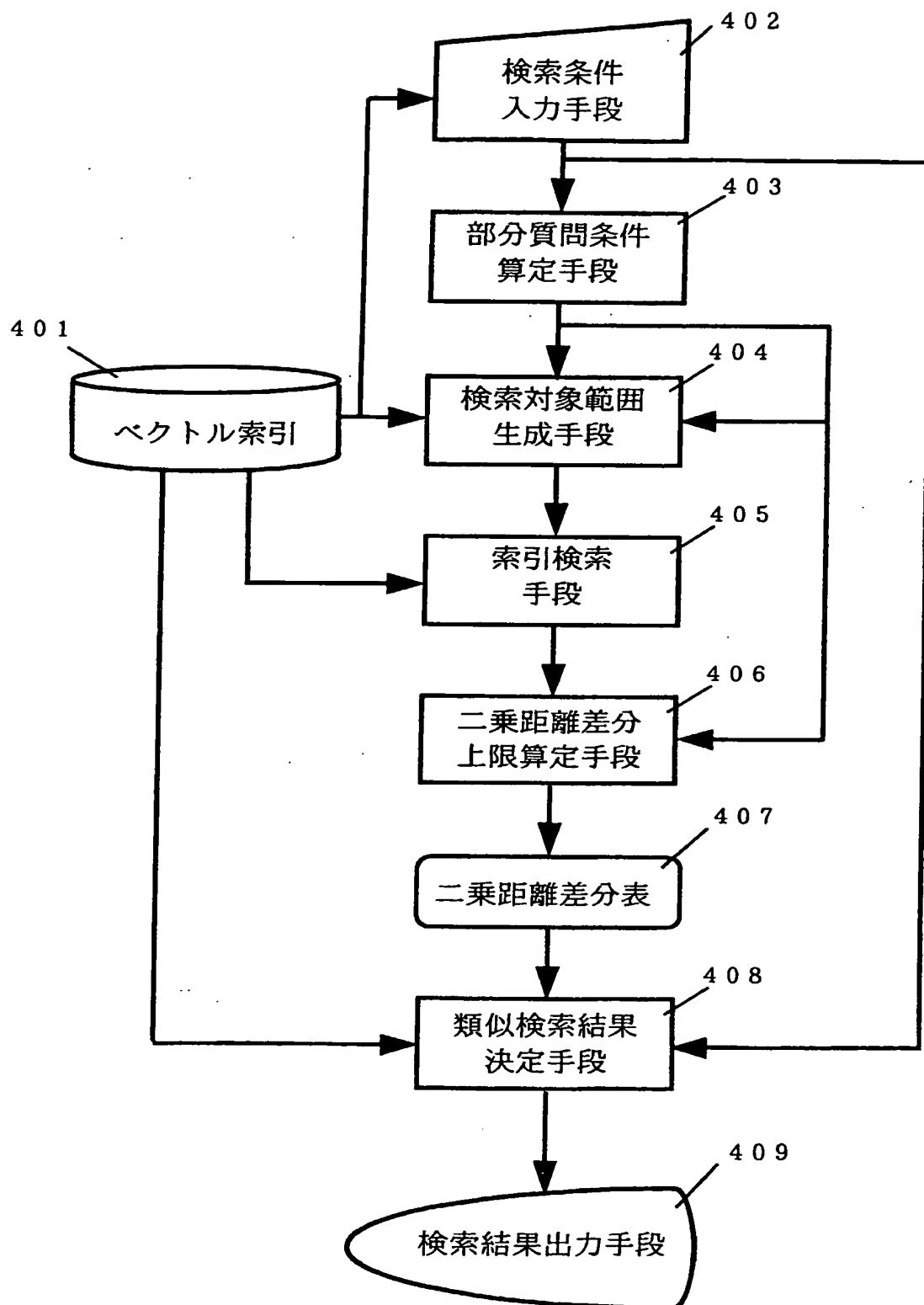
FIG. 3



THIS PAGE BLANK (USPTO)

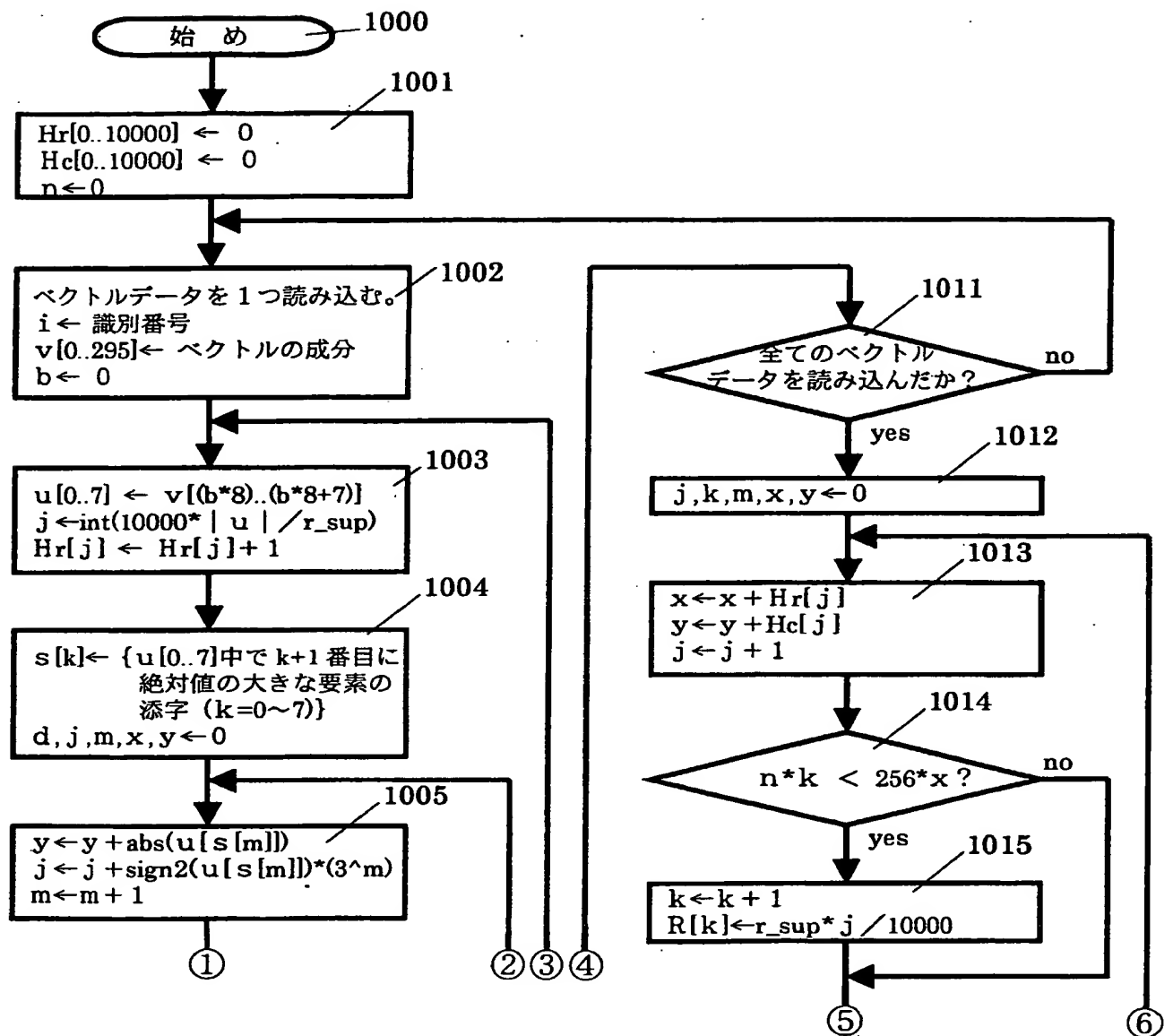
4/29

FIG. 4



THIS PAGE BLANK (USPTO)

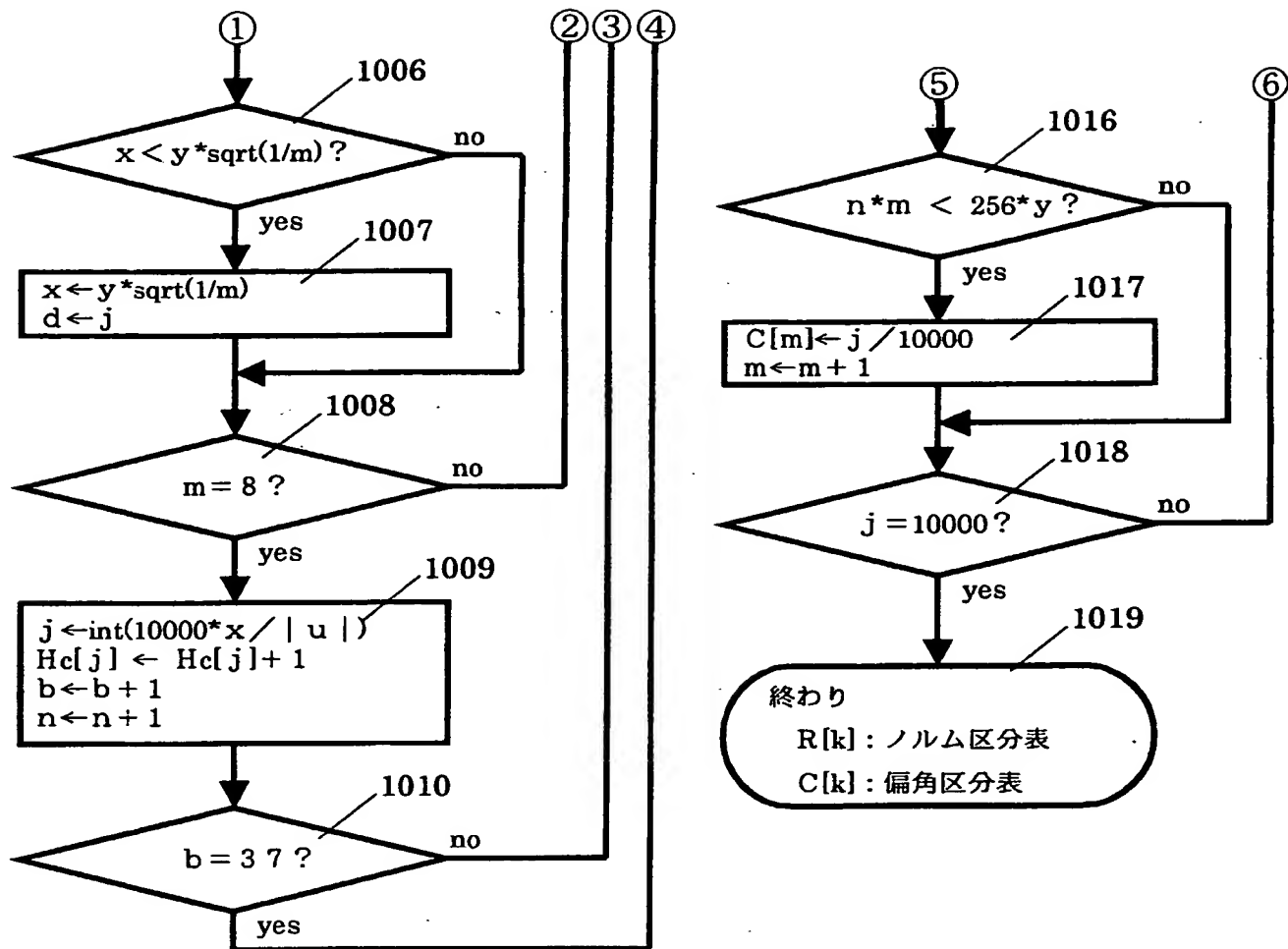
FIG. 5A



THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

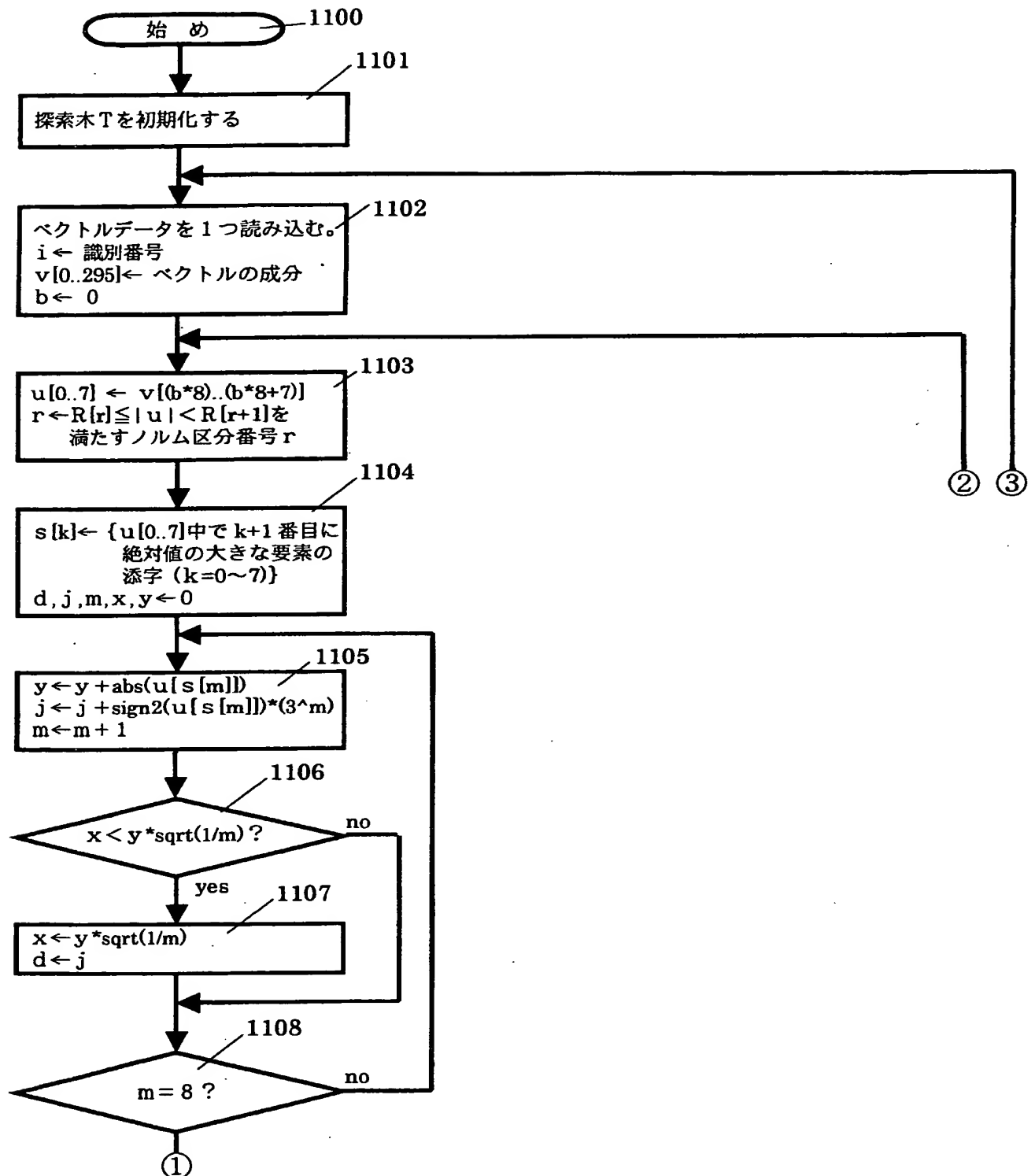
FIG. 5B



THIS PAGE BLANK (USPTO)

7/29

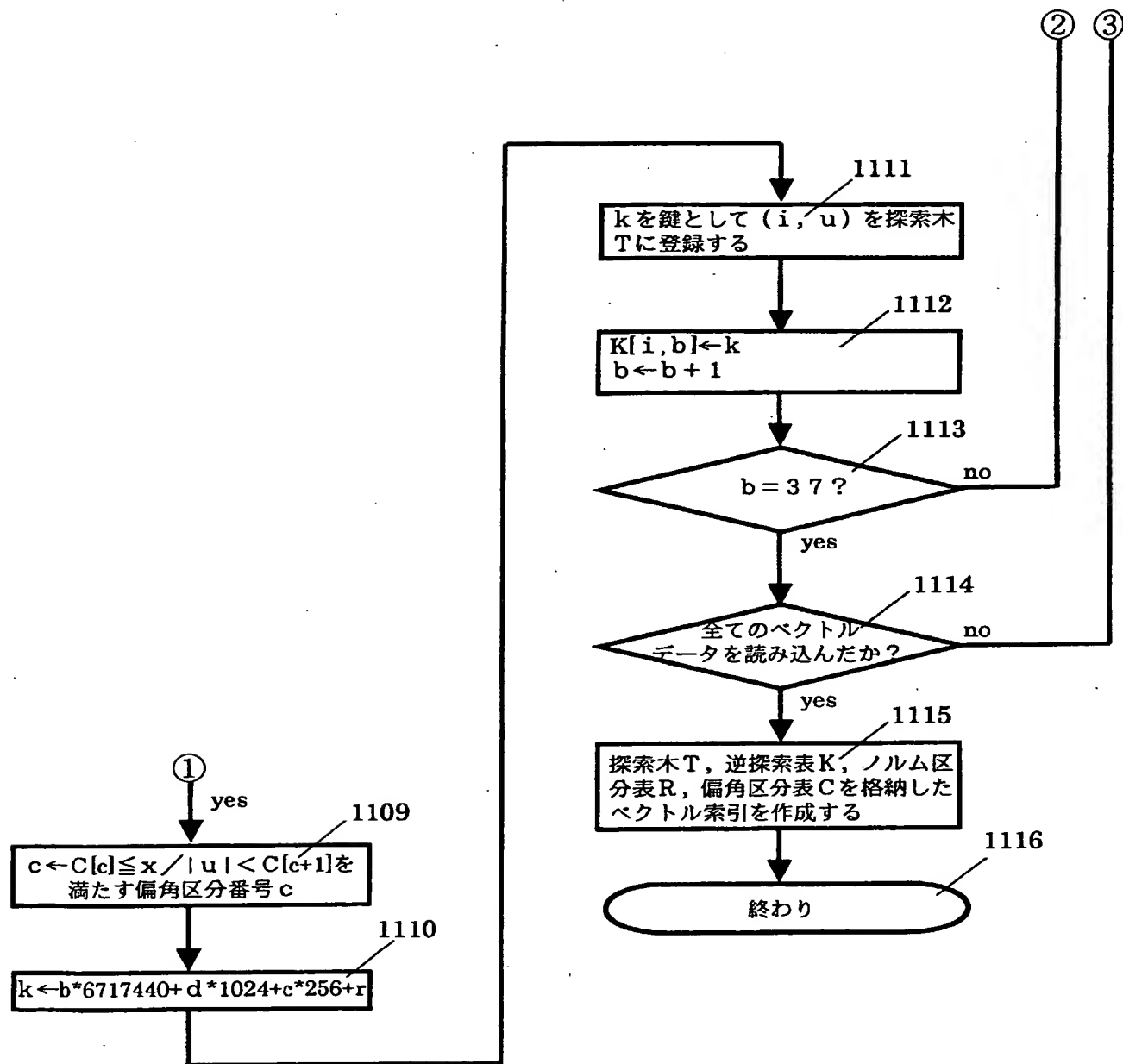
FIG. 6A



THIS PAGE BLANK (USPTO)

8/29

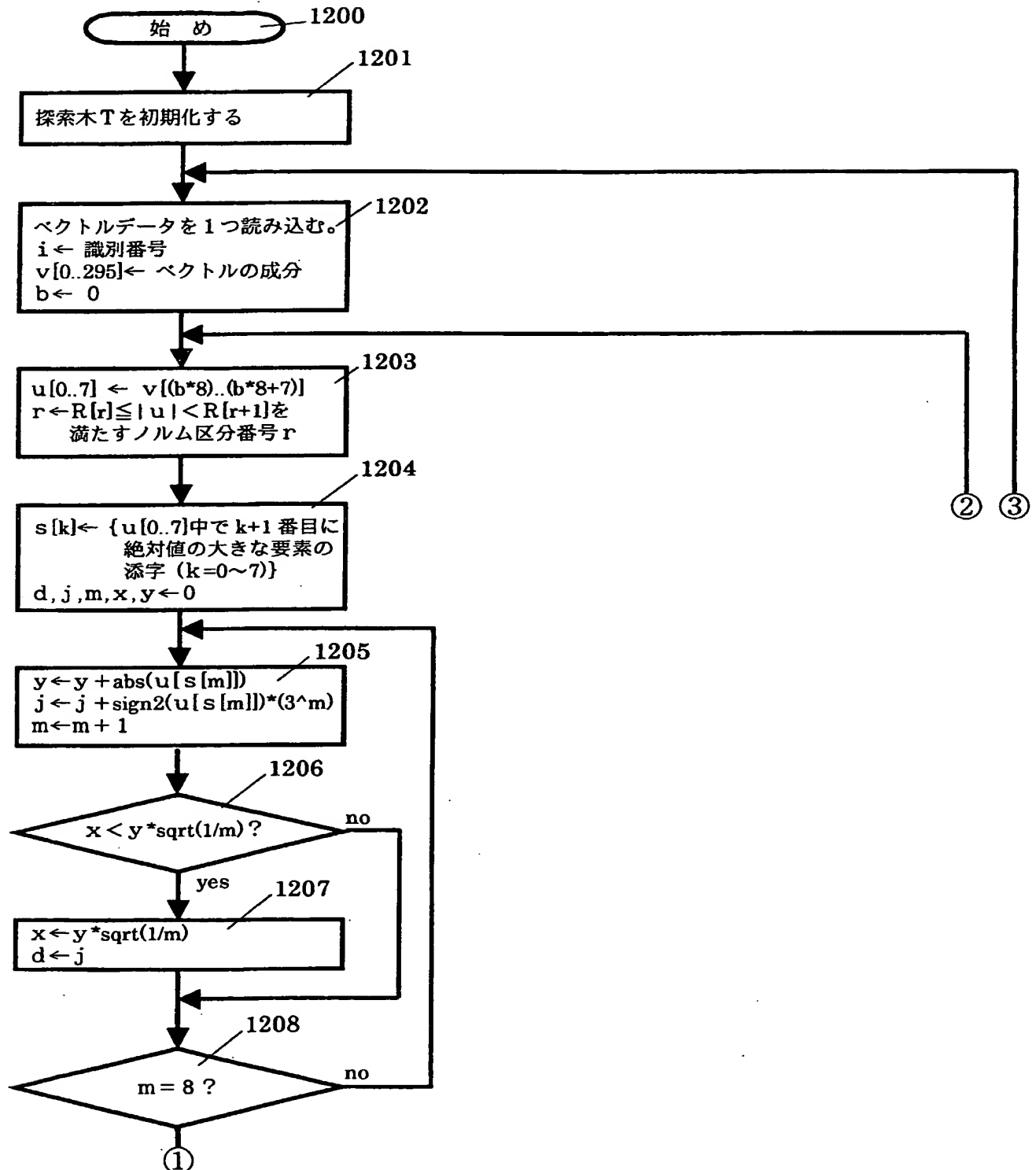
FIG. 6B



THIS PAGE BLANK (USPTO)

9/29

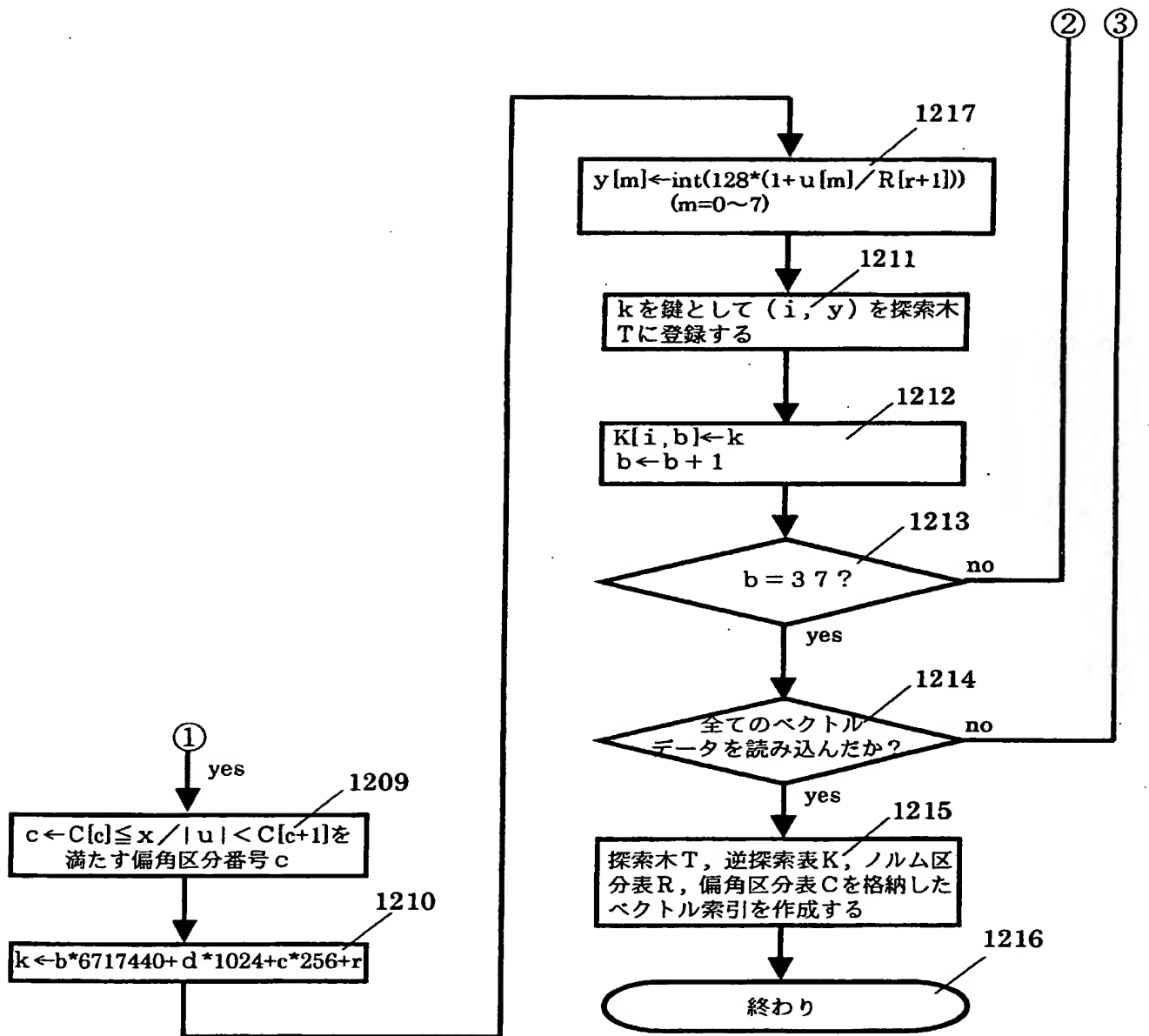
FIG. 7A



THIS PAGE BLANK (USPTO)

10/29

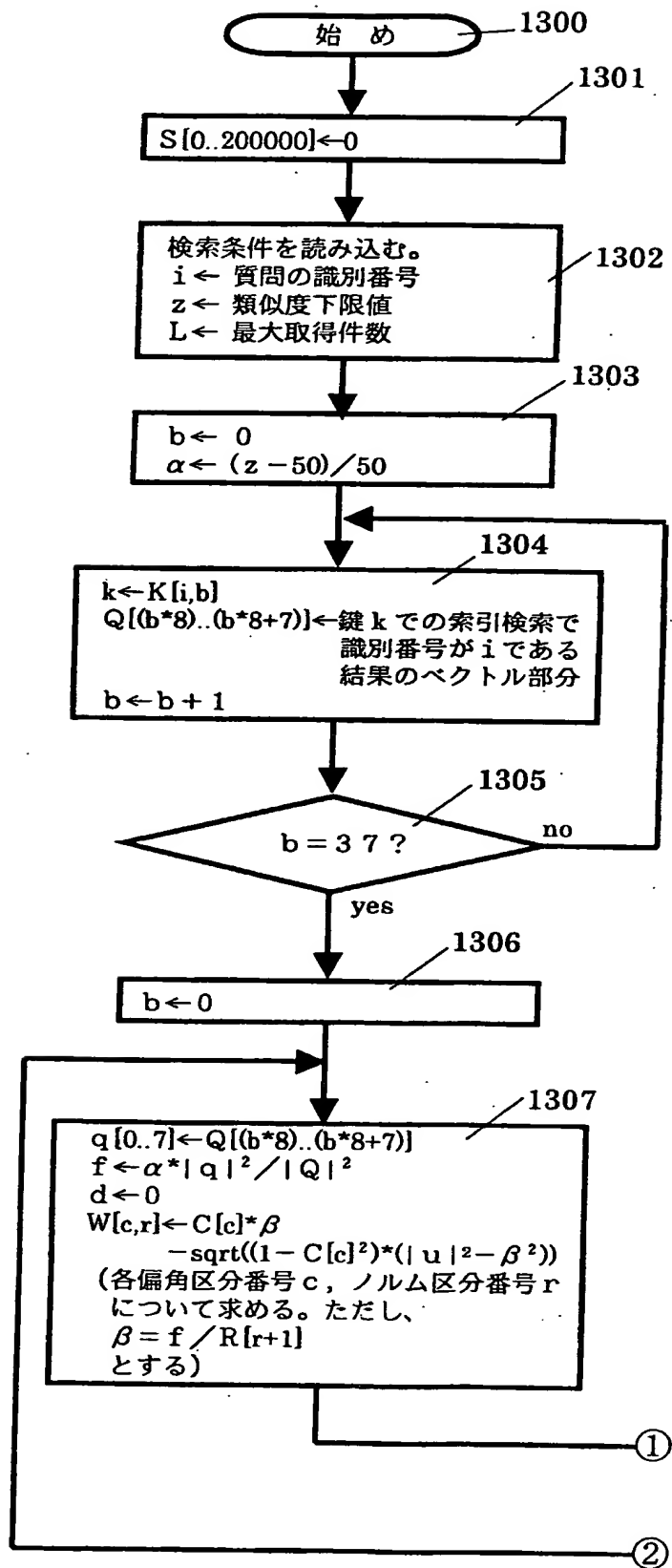
FIG. 7B



THIS PAGE BLANK (USPTO)

11/29

FIG. 8A

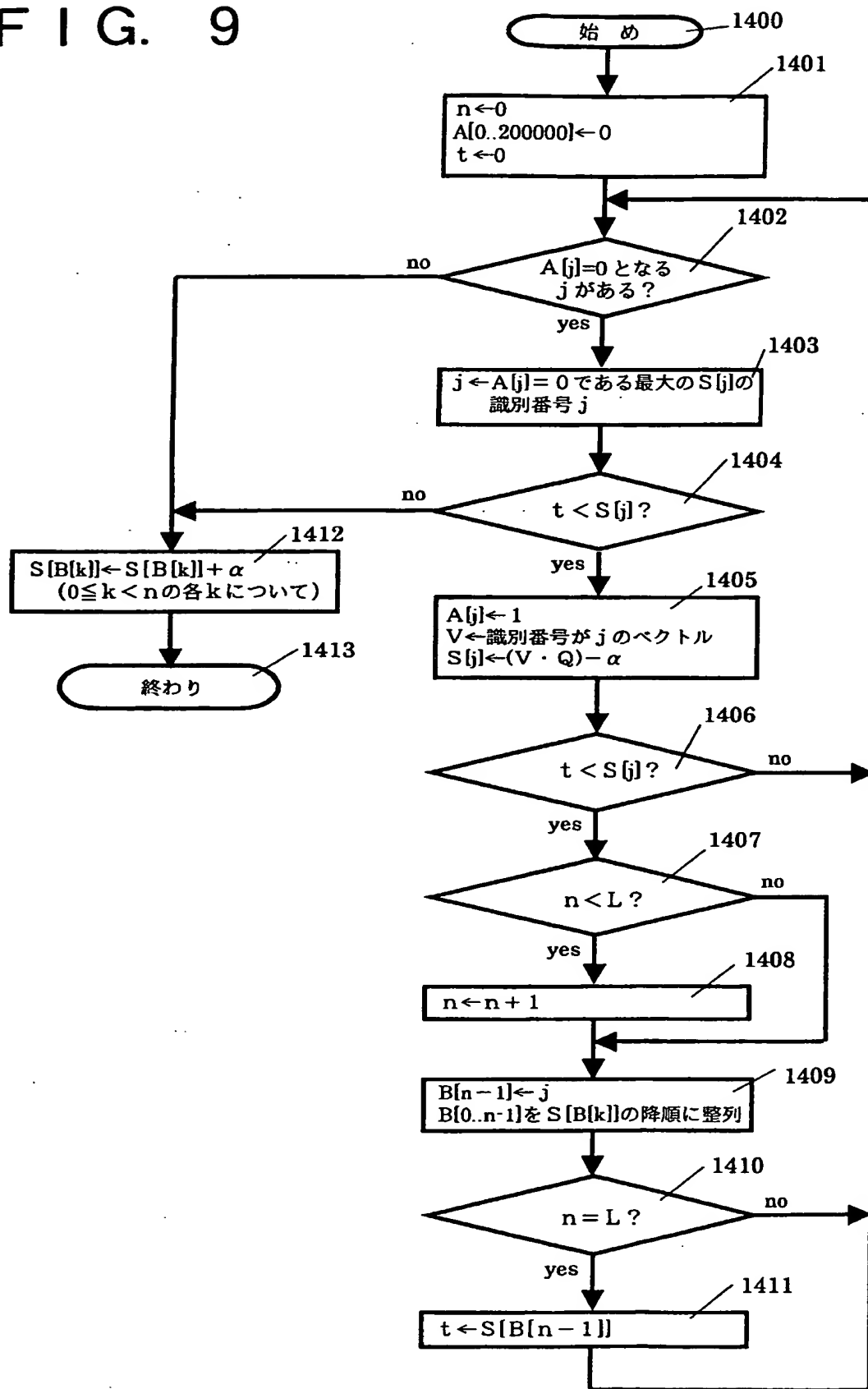


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

13/29

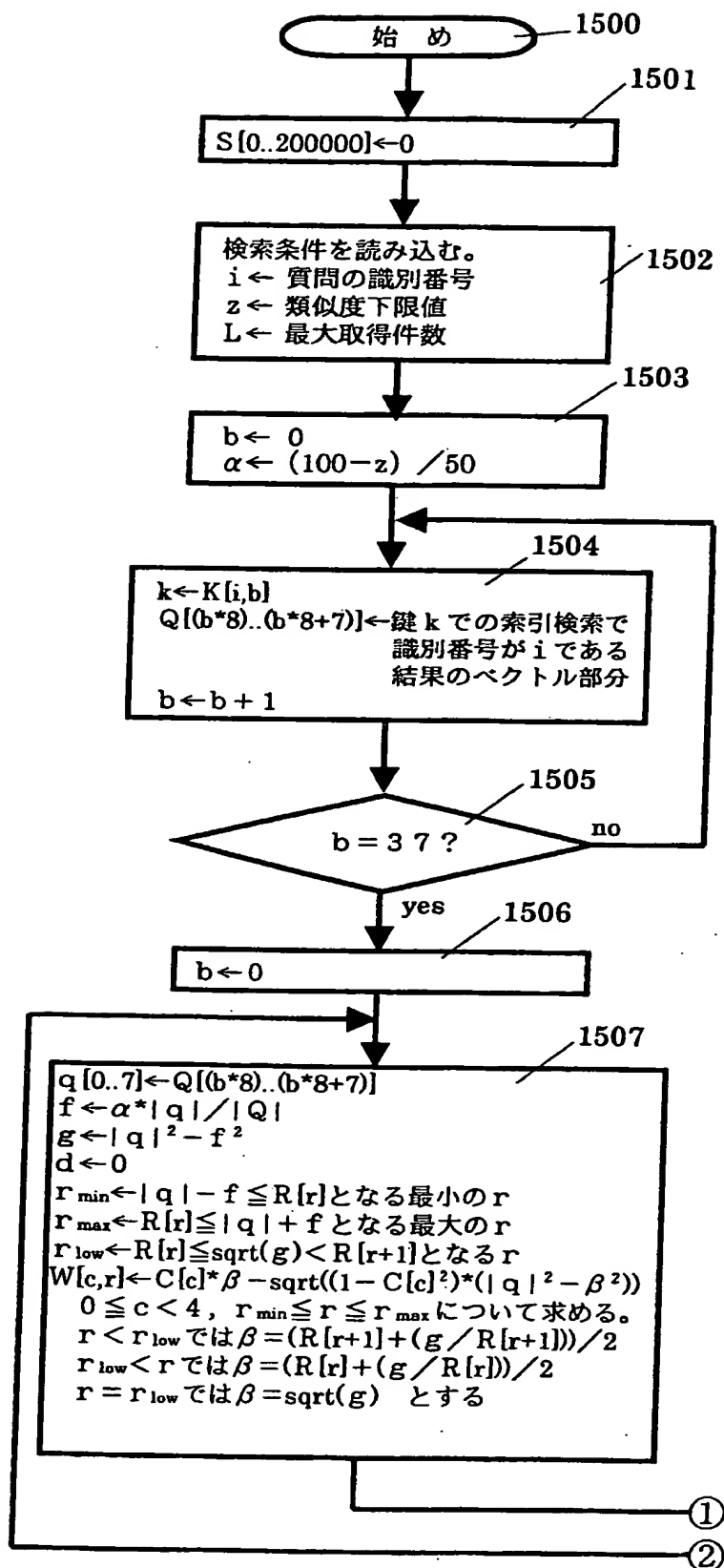
FIG. 9



THIS PAGE BLANK (USPTO)

14/29

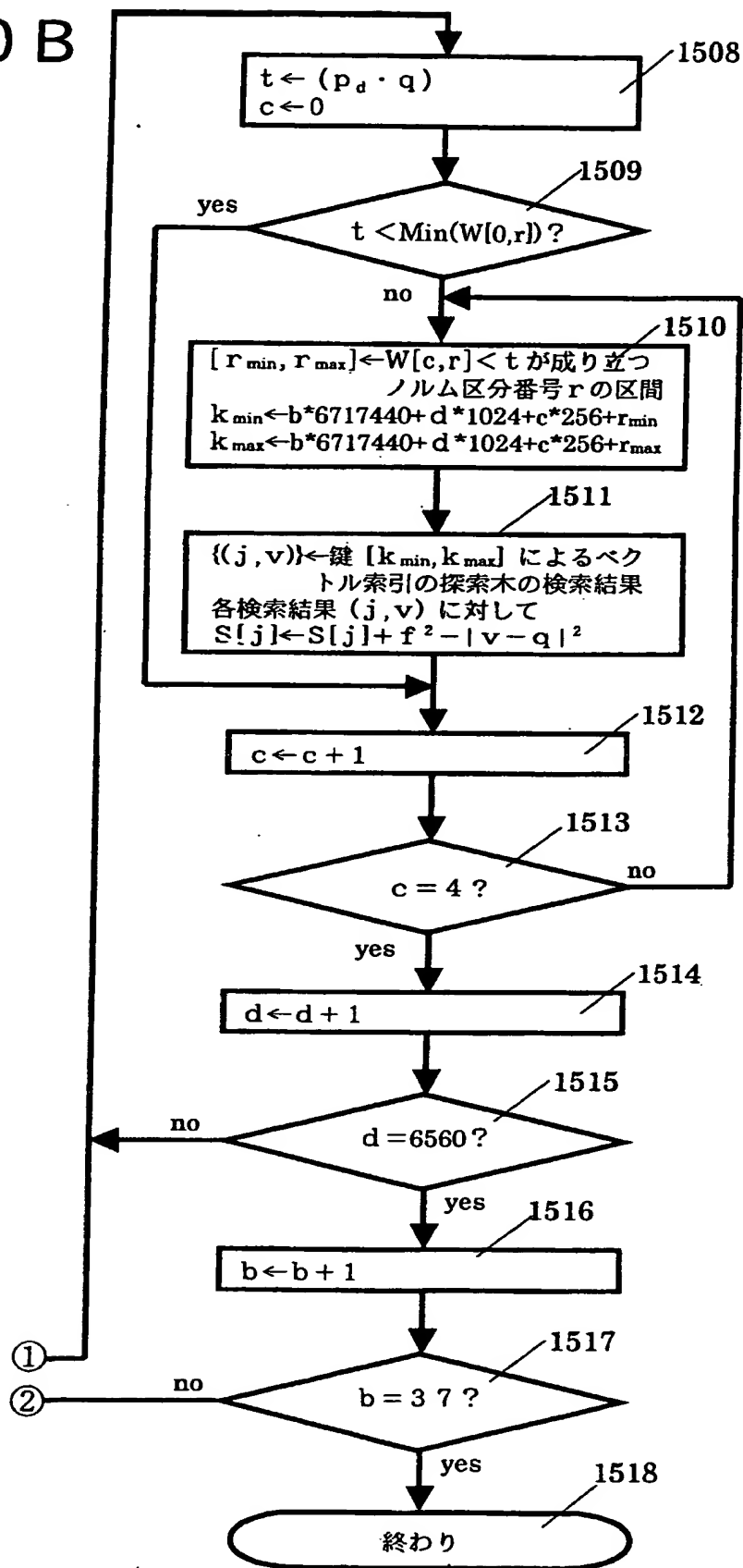
FIG. 10A



THIS PAGE BLANK (USPTO)

15/29

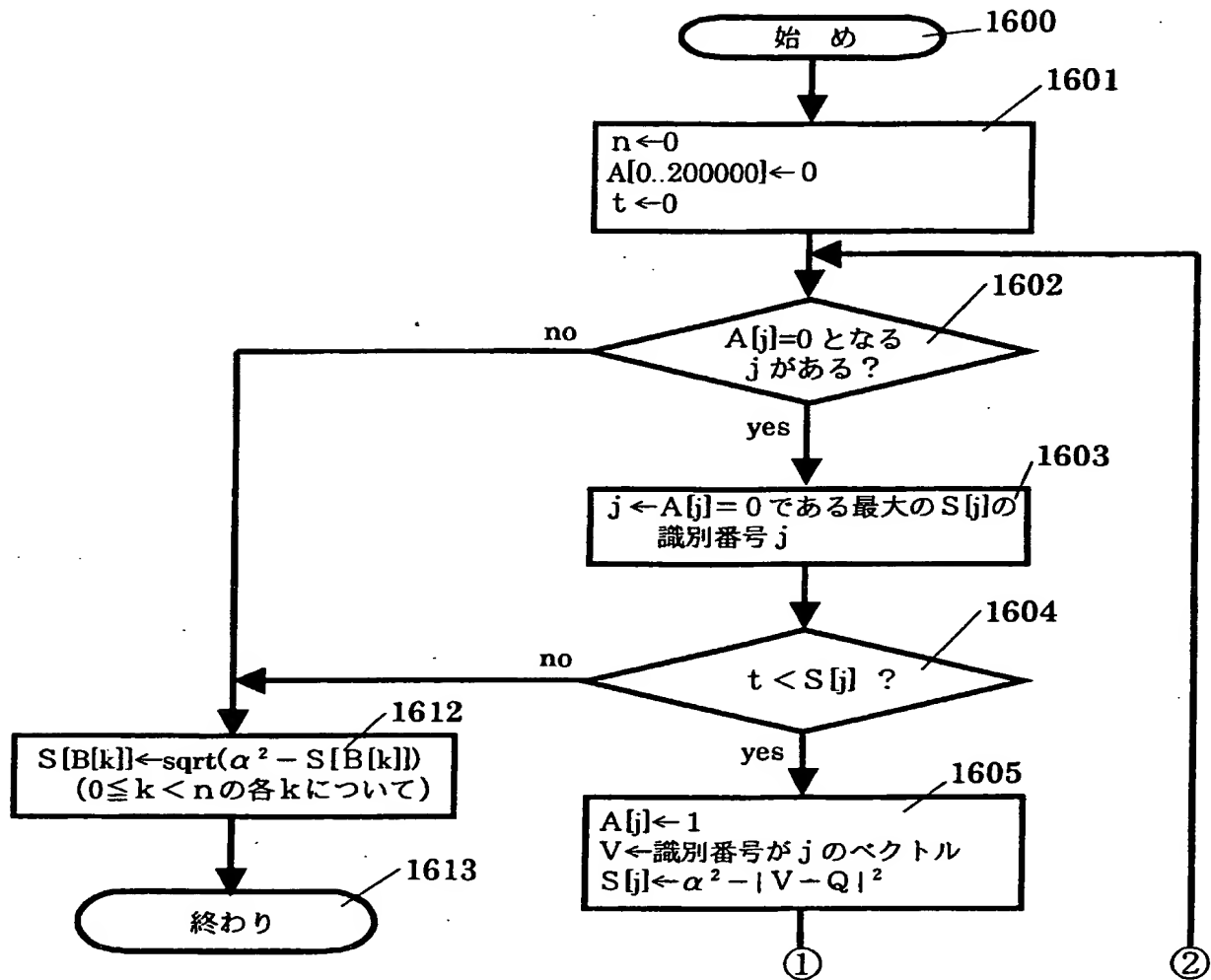
FIG. 10B



THIS PAGE BLANK (USPTO)

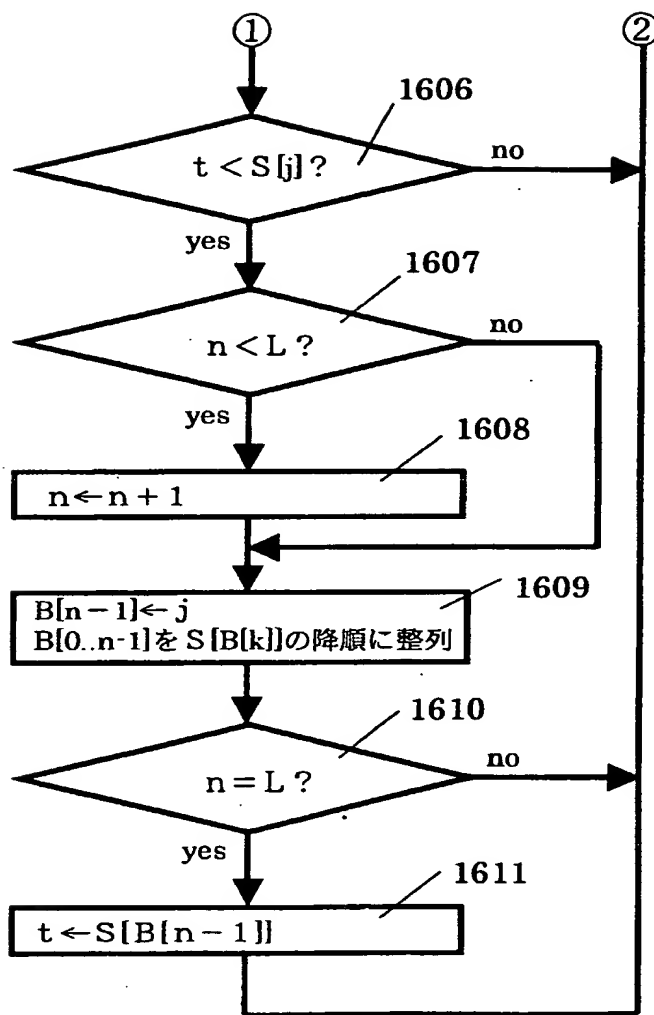
16/29

FIG. 11A



THIS PAGE BLANK (USPTO)

FIG. 11B



THIS PAGE BLANK (USPTO)

18/29

FIG. 12A

```

1 +0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.007255 -0.007715 -0.025648 +0.016061
  -0.060584 -0.013593 -0.020985 -0.112403 -0.012045 +0.044741 +0.026761 +0.078339 +0.048166 +0.043434
  +0.100093 +0.009913 +0.085770 +0.101257 +0.072163 -0.066112 +0.059376 -0.020159 +0.051960 -0.129138
  -0.028065 +0.027535 +0.028316 +0.050490 +0.015931 -0.040316 -0.013109 -0.014728 -0.004639 -0.021525
  -0.000471 -0.033506 +0.013866 -0.054646 +0.067350 +0.042063 +0.041963 -0.006444 -0.092581 +0.004488
  +0.004741 +0.009351 +0.038429 -0.042254 -0.027641 -0.068727 +0.037185 -0.003393 -0.040649 +0.013169
  +0.020619 +0.025594 -0.019990 -0.117804 +0.005791 -0.027860 +0.000220 -0.038765 -0.029964 +0.020038
  +0.032435 -0.027518 -0.063942 +0.085381 +0.038776 +0.051395 +0.004047 +0.092011 -0.076222 +0.096729
  -0.018331 +0.115754 -0.038478 +0.131147 -0.074560 +0.080634 -0.186932 +0.024004 +0.047046 -0.075571
  +0.121789 -0.055221 -0.001166 -0.053469 -0.086326 +0.011837 -0.060801 +0.222437 -0.055550 -0.117881
  -0.020700 -0.028172 -0.121642 -0.160389 +0.147645 -0.037681 -0.057998 +0.104025 +0.251415 -0.029438
  +0.030504 -0.048312 -0.072984 -0.088780 +0.041684 +0.127138 +0.061804 +0.064147 -0.016586 +0.024305
  +0.060558 -0.004070 +0.094040 -0.011500 +0.000545 +0.083231 +0.016565 +0.081034 +0.073438 -0.006857
  -0.008995 +0.023537 +0.066849 -0.035310 +0.005572 -0.015236 +0.109983 -0.185597 +0.016643 +0.032632
  -0.075726 -0.110307 +0.038577 +0.038475 -0.042287 +0.082878 +0.035997 -0.009888 +0.081286 +0.063583
  -0.041429 +0.025969 -0.040406 +0.005639 +0.032087 +0.007947 +0.041689 +0.040077 +0.067726 -0.101670
  -0.091183 +0.167914 -0.080320 +0.049351 +0.069409 +0.063139 -0.038358 -0.126212 +0.058109 +0.031847
  -0.014998 -0.022995 +0.054876 +0.033124 -0.065283 -0.058574 +0.049729 -0.046552 +0.042485 -0.006179
  -0.058764 +0.079383 +0.000817 -0.001482 -0.036410 -0.036097 -0.045920 -0.001729 +0.039971 +0.083165
  -0.023112 +0.014492 +0.028403 +0.047480 +0.038502 +0.028348 +0.055128 +0.045340 -0.066148 +0.018156
  -0.008535 -0.042836 +0.006119 -0.037691 +0.018055 +0.035741 -0.023394 +0.012401 -0.070880 +0.010066
  -0.013264 -0.031192 -0.064061 -0.026757 -0.028246 +0.078634 +0.013295 +0.011129 +0.028807 +0.012339
  +0.007173 -0.008856 +0.040397 +0.039853 +0.085247 -0.053102 +0.052307 +0.065223 +0.116747 +0.013464
  -0.004875 +0.019186 +0.024114 -0.056101 -0.024008 +0.061251 -0.043466 -0.017640 +0.081001 -0.014824
  -0.003836 +0.059081 +0.051690 -0.032798 +0.039059 +0.020370 +0.015096 +0.051693 +0.015507 -0.041601
  -0.000192 -0.065087 +0.018487 -0.040415 +0.036173 -0.011809 +0.010862 +0.005944 +0.028534 -0.031335
  +0.023075 +0.033037 +0.063589 +0.014185 +0.006539 +0.002593 -0.023986 -0.038277 -0.009555 -0.018987
  +0.052526 +0.035448 +0.013042 +0.023662 +0.011775 -0.055742 -0.008120 -0.040546 -0.023508 -0.069309
  +0.037886 +0.041494 -0.038487 -0.035241 +0.020432 -0.008060 +0.002984 +0.070241 +0.069379 +0.020206
  +0.032996 +0.047815 +0.046106 +0.001794 +0.035342 -0.003895

```

THIS PAGE BLANK (USPTO)

19/29

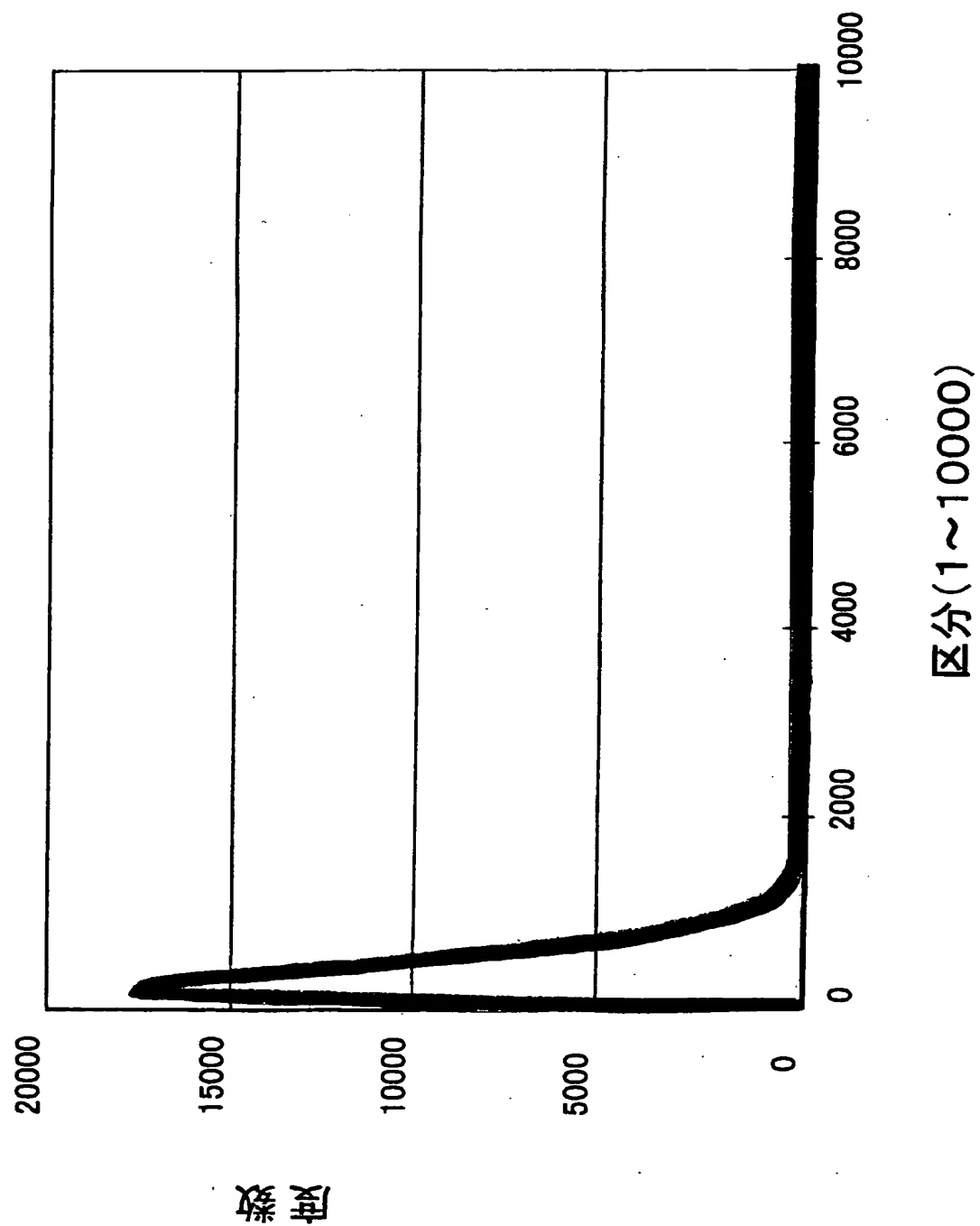
FIG. 12B

2 +0.028972 -0.012757 -0.015597 +0.019727 +0.009386 -0.016593 +0.003627 +0.006288 -0.019184 +0.020306
 -0.057163 -0.017815 -0.026345 -0.102036 +0.002587 +0.037785 +0.029168 +0.076061 +0.043901 +0.040040
 +0.123462 +0.001139 +0.085437 +0.108889 +0.052652 -0.048914 +0.060612 -0.005019 +0.030421 -0.153633
 -0.041444 +0.038908 +0.006823 +0.069954 +0.028216 -0.043207 -0.030092 +0.013753 +0.023770 -0.018313
 +0.008825 -0.036443 +0.001076 -0.067721 +0.046034 +0.030717 +0.017880 -0.036936 -0.093124 -0.000883
 +0.027865 +0.007906 +0.005978 -0.024367 -0.012682 -0.054200 +0.025934 +0.000926 -0.047710 +0.009757
 +0.005940 +0.020855 -0.070890 -0.113381 +0.004988 -0.038150 -0.000456 -0.024836 -0.007560 +0.025912
 +0.009004 -0.053047 -0.079142 +0.085440 +0.027876 +0.051104 +0.016944 +0.082277 -0.071359 +0.107308
 +0.005041 +0.112418 -0.009138 +0.119492 -0.069016 +0.123591 -0.166734 +0.032382 +0.005430 -0.030192
 +0.116327 -0.077304 +0.003280 -0.006984 -0.055858 +0.022018 -0.110375 +0.197565 -0.038060 -0.085170
 -0.065823 -0.021350 -0.104387 -0.147696 +0.111377 -0.028678 -0.097095 +0.064212 +0.255376 -0.011000
 +0.020901 -0.032671 -0.092765 -0.063843 +0.008917 +0.106446 +0.070094 +0.078741 -0.012886 -0.003581
 +0.069363 +0.021164 +0.046900 -0.021002 -0.008879 +0.052981 +0.006370 +0.081378 +0.054328 -0.006424
 -0.006277 +0.013635 +0.117156 -0.037470 +0.014036 -0.048765 +0.093100 -0.147319 +0.028556 -0.017833
 -0.070005 -0.123845 +0.013978 +0.006964 -0.047420 +0.100905 -0.019278 -0.009641 +0.057287 +0.058665
 -0.063796 +0.019097 -0.045014 -0.036129 +0.022014 +0.071405 +0.028573 +0.046653 +0.063911 -0.048555
 -0.070203 +0.205558 -0.051782 +0.102727 +0.042066 +0.028359 -0.021939 -0.082837 +0.064817 +0.017215
 -0.042670 -0.031901 +0.037475 +0.055012 -0.012237 -0.067371 +0.072587 -0.009949 +0.053991 +0.019722
 -0.035742 +0.081726 +0.019732 +0.013624 -0.031871 -0.009025 -0.064237 +0.002162 +0.014326 +0.103617
 -0.048376 +0.034422 -0.019797 +0.041018 +0.089878 +0.072000 +0.030657 +0.040709 -0.071603 +0.005629

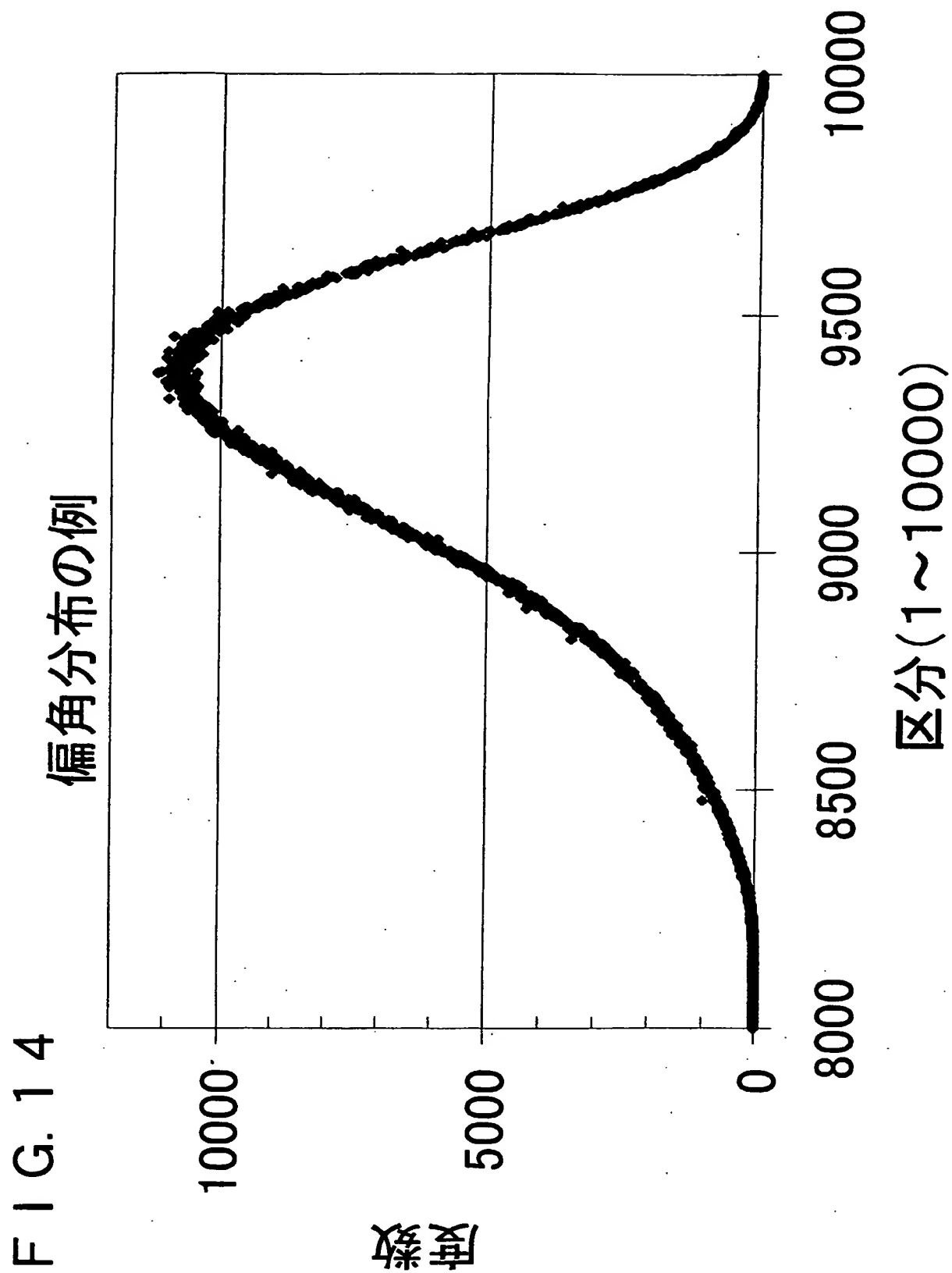
THIS PAGE BLANK (USPTO)

FIG. 13

ノルム分布の例



THIS PAGE BLANK (USPTO)



THIS PAGE BLANK (USPTO)

FIG. 15A

ノルム区分表の例 (2 5 6 区分)

	0	1	2	3	4	5	6	7	8	9
0	0.00000	0.03320	0.04112	0.04617	0.04999	0.05354	0.05687	0.05940	0.06182	0.06416
10	0.06585	0.06805	0.06965	0.07121	0.07275	0.07425	0.07572	0.07716	0.07858	0.07951
20	0.08088	0.08178	0.08312	0.08400	0.08530	0.08616	0.08701	0.08827	0.08910	0.08992
30	0.09073	0.09194	0.09273	0.09352	0.09431	0.09508	0.09585	0.09662	0.09737	0.09813
40	0.09887	0.09961	0.10035	0.10108	0.10180	0.10252	0.10324	0.10395	0.10465	0.10535
50	0.10605	0.10639	0.10708	0.10776	0.10844	0.10912	0.10979	0.11046	0.11079	0.11145
60	0.11211	0.11276	0.11309	0.11374	0.11438	0.11502	0.11566	0.11598	0.11661	0.11724
70	0.11786	0.11849	0.11879	0.11941	0.12003	0.12064	0.12094	0.12155	0.12215	0.12275
80	0.12305	0.12365	0.12424	0.12483	0.12542	0.12571	0.12629	0.12687	0.12745	0.12774
90	0.12831	0.12888	0.12945	0.13002	0.13030	0.13087	0.13143	0.13198	0.13254	0.13282
100	0.13337	0.13392	0.13447	0.13501	0.13556	0.13583	0.13637	0.13690	0.13744	0.13797
110	0.13851	0.13904	0.13956	0.13983	0.14035	0.14087	0.14139	0.14191	0.14243	0.14295
120	0.14346	0.14397	0.14448	0.14499	0.14549	0.14600	0.14650	0.14700	0.14750	0.14800

THIS PAGE BLANK (USPTO)

FIG. 15B

130	0.14849	0.14899	0.14948	0.14997	0.15046	0.15095	0.15143	0.15192	0.15240	0.15288
140	0.15336	0.15408	0.15456	0.15503	0.15551	0.15598	0.15645	0.15715	0.15762	0.15808
150	0.15878	0.15924	0.15970	0.16016	0.16085	0.16131	0.16199	0.16244	0.16289	0.16357
160	0.16402	0.16469	0.16513	0.16580	0.16624	0.16690	0.16734	0.16800	0.16866	0.16909
170	0.16974	0.17039	0.17104	0.17147	0.17211	0.17275	0.17338	0.17402	0.17465	0.17507
180	0.17570	0.17633	0.17716	0.17778	0.17840	0.17902	0.17963	0.18024	0.18106	0.18166
190	0.18227	0.18308	0.18368	0.18447	0.18507	0.18586	0.18665	0.18724	0.18803	0.18881
200	0.18958	0.19036	0.19113	0.19190	0.19266	0.19342	0.19437	0.19512	0.19606	0.19681
210	0.19774	0.19867	0.19959	0.20051	0.20143	0.20252	0.20342	0.20450	0.20540	0.20647
220	0.20754	0.20860	0.20983	0.21087	0.21209	0.21330	0.21450	0.21587	0.21706	0.21858
230	0.21992	0.22142	0.22291	0.22438	0.22602	0.22780	0.22957	0.23148	0.23338	0.23557
240	0.23774	0.24005	0.24249	0.24520	0.24818	0.25142	0.25505	0.25919	0.26369	0.26921
250	0.27595	0.28434	0.29600	0.31512	0.35936	0.49100	0.85733			

THIS PAGE BLANK (USPTO)

FIG. 16

偏角区分表の例（4区分）

区分番号	区分境界
0	0.8274
1	0.9079
2	0.9301
3	0.9486
4	1.0000

THIS PAGE BLANK (USPTO)

25/29

FIG. 17A

区分	0	1	2	3	4	5	6	7	8	9
b=0, c=0										
0	+9.99999	+0.03142	+0.01968	+0.01436	+0.01008	+0.00750	+0.00536	+0.00355	+0.00220	+0.00100
10	+0.00010	-0.00089	-0.00164	-0.00233	-0.00298	-0.00347	-0.00404	-0.00447	-0.00498	-0.00537
20	-0.00574	-0.00609	-0.00643	-0.00675	-0.00706	-0.00736	-0.00764	-0.00792	-0.00812	-0.00838
30	-0.00857	-0.00881	-0.00904	-0.00921	-0.00938	-0.00959	-0.00975	-0.00996	-0.01010	-0.01025
40	-0.01039	-0.01058	-0.01071	-0.01084	-0.01097	-0.01110	-0.01122	-0.01139	-0.01150	-0.01162
50	-0.01173	-0.01185	-0.01196	-0.01206	-0.01217	-0.01227	-0.01237	-0.01247	-0.01257	-0.01267
60	-0.01276	-0.01286	-0.01292	-0.01301	-0.01310	-0.01319	-0.01327	-0.01336	-0.01344	-0.01352
70	-0.01360	-0.01368	-0.01376	-0.01381	-0.01389	-0.01397	-0.01404	-0.01411	-0.01418	-0.01426
80	-0.01433	-0.01439	-0.01446	-0.01453	-0.01457	-0.01464	-0.01470	-0.01477	-0.01483	-0.01489
90	-0.01496	-0.01502	-0.01508	-0.01514	-0.01519	-0.01525	-0.01531	-0.01536	-0.01542	-0.01547
100	-0.01553	-0.01558	-0.01564	-0.01569	-0.01574	-0.01579	-0.01584	-0.01589	-0.01594	-0.01599
110	-0.01604	-0.01609	-0.01613	-0.01618	-0.01623	-0.01627	-0.01633	-0.01638	-0.01642	-0.01646
120	-0.01651	-0.01656	-0.01661	-0.01665	-0.01669	-0.01673	-0.01678	-0.01682	-0.01686	-0.01692
130	-0.01695	-0.01699	-0.01704	-0.01708	-0.01712	-0.01717	-0.01720	-0.01725	-0.01729	-0.01733
140	-0.01737	-0.01741	-0.01745	-0.01749	-0.01753	-0.01757	-0.01761	-0.01766	-0.01769	-0.01773
150	-0.01777	-0.01780	-0.01784	-0.01788	-0.01792	-0.01796	-0.01800	-0.01804	-0.01808	-0.01812
160	-0.01816	-0.01819	-0.01823	-0.01827	-0.01830	-0.01835	-0.01838	-0.01842	-0.01846	-0.01849
170	-0.01854	-0.01857	-0.01861	-0.01864	-0.01868	-0.01872	-0.01875	-0.01879	-0.01883	-0.01887
180	-0.01891	-0.01894	-0.01898	-0.01902	-0.01906	-0.01909	-0.01913	-0.01917	-0.01920	-0.01924
190	-0.01928	-0.01932	-0.01936	-0.01939	-0.01943	-0.01947	-0.01951	-0.01955	-0.01958	-0.01962
200	-0.01966	-0.01970	-0.01974	-0.01978	-0.01982	-0.01986	-0.01990	-0.01994	-0.01998	-0.02002
210	-0.02006	-0.02010	-0.02015	-0.02019	-0.02023	-0.02027	-0.02032	-0.02036	-0.02040	-0.02045
220	-0.02049	-0.02054	-0.02059	-0.02063	-0.02068	-0.02073	-0.02078	-0.02083	-0.02088	-0.02094
230	-0.02099	-0.02104	-0.02110	-0.02116	-0.02121	-0.02128	-0.02134	-0.02140	-0.02147	-0.02154
240	-0.02161	-0.02169	-0.02177	-0.02185	-0.02194	-0.02204	-0.02215	-0.02227	-0.02241	-0.02256
250	-0.02275	-0.02299	-0.02334	-0.02401	-0.02527					

THIS PAGE BLANK (USPTO)

FIG. 17B

b=0, c=3

0	+9.99999	+0.04126	+0.03220	+0.02771	+0.02395	+0.02162	+0.01966	+0.01798	+0.01672	+0.01559
10	+0.01473	+0.01378	+0.01306	+0.01239	+0.01176	+0.01129	+0.01073	+0.01030	+0.00980	+0.00942
20	+0.00905	+0.00871	+0.00837	+0.00805	+0.00774	+0.00744	+0.00715	+0.00688	+0.00668	+0.00642
30	+0.00623	+0.00598	+0.00575	+0.00558	+0.00541	+0.00519	+0.00503	+0.00482	+0.00467	+0.00452
40	+0.00438	+0.00419	+0.00405	+0.00392	+0.00378	+0.00365	+0.00353	+0.00336	+0.00324	+0.00312
50	+0.00300	+0.00289	+0.00278	+0.00267	+0.00256	+0.00245	+0.00235	+0.00224	+0.00214	+0.00204
60	+0.00194	+0.00185	+0.00178	+0.00169	+0.00160	+0.00151	+0.00142	+0.00133	+0.00124	+0.00116
70	+0.00107	+0.00099	+0.00091	+0.00085	+0.00077	+0.00070	+0.00062	+0.00054	+0.00047	+0.00039
80	+0.00032	+0.00025	+0.00018	+0.00010	+0.00006	-0.00001	-0.00008	-0.00015	-0.00021	-0.00028
90	-0.00034	-0.00040	-0.00047	-0.00053	-0.00059	-0.00065	-0.00071	-0.00077	-0.00083	-0.00089
100	-0.00094	-0.00100	-0.00106	-0.00111	-0.00117	-0.00122	-0.00127	-0.00133	-0.00138	-0.00143
110	-0.00148	-0.00153	-0.00158	-0.00163	-0.00168	-0.00173	-0.00179	-0.00184	-0.00189	-0.00193
120	-0.00198	-0.00204	-0.00208	-0.00213	-0.00217	-0.00221	-0.00227	-0.00231	-0.00236	-0.00241
130	-0.00245	-0.00249	-0.00255	-0.00259	-0.00263	-0.00268	-0.00272	-0.00277	-0.00281	-0.00286
140	-0.00289	-0.00294	-0.00298	-0.00303	-0.00306	-0.00311	-0.00316	-0.00320	-0.00324	-0.00328
150	-0.00333	-0.00336	-0.00340	-0.00345	-0.00349	-0.00353	-0.00357	-0.00361	-0.00366	-0.00370
160	-0.00374	-0.00378	-0.00382	-0.00386	-0.00389	-0.00394	-0.00398	-0.00402	-0.00406	-0.00410
170	-0.00414	-0.00418	-0.00423	-0.00426	-0.00430	-0.00434	-0.00438	-0.00442	-0.00446	-0.00451
180	-0.00455	-0.00458	-0.00463	-0.00467	-0.00470	-0.00474	-0.00478	-0.00483	-0.00486	-0.00491
190	-0.00494	-0.00499	-0.00503	-0.00507	-0.00511	-0.00515	-0.00519	-0.00523	-0.00527	-0.00532
200	-0.00536	-0.00540	-0.00544	-0.00548	-0.00553	-0.00557	-0.00562	-0.00566	-0.00571	-0.00575
210	-0.00579	-0.00584	-0.00589	-0.00593	-0.00598	-0.00602	-0.00607	-0.00612	-0.00617	-0.00622
220	-0.00627	-0.00632	-0.00637	-0.00642	-0.00647	-0.00653	-0.00658	-0.00663	-0.00669	-0.00675
230	-0.00680	-0.00686	-0.00692	-0.00699	-0.00705	-0.00712	-0.00719	-0.00726	-0.00733	-0.00741
240	-0.00749	-0.00757	-0.00766	-0.00775	-0.00786	-0.00797	-0.00808	-0.00821	-0.00837	-0.00854
250	-0.00875	-0.00901	-0.00941	-0.01015	-0.01157					

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

FIG. 18B

$$b=1, c=1$$

THIS PAGE BLANK (USPTO)

FIG. 18C

$$b=1, c=2$$
[illegible]

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/09079

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G06F17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2001
 Kokai Jitsuyo Shinan Koho 1971-2001 Jitsuyo Shinan Toroku Koho 1996-2001

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
 JICST FILE (JOIS) (vector, norm, index, high dimension, database) (in Japanese)
 WPI, INSPEC (DIALOG) (vector, near neighbor, index, search, retrieve, high dimension, database)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LEE J.Y. et. al., "A fast retrieval method for image features.", TENCON 99. Proceedings of the IEEE Region 10 Conference, Vol.1, 15 September, 1999 (15.09.99), pp.577-580	1-29
A	NENE S.A. & NAYAR S.K. "A simple algorithm for nearest neighbor search in high dimensions," IEEE Transactions on Pattern Analysis and Intelligence, Vol.19, No.9, September, 1997 (09.97), pp.989-1003	1-29
A	Washizawa, "k-Nearest Neighbor Tansaku no Kousoku Keisan Algorithm", Gazou Denshi Gakkaishi, Vol.27, No.5, (Japan), 25 October, 1998 (25.10.98), pp.513-520	1-29
A	KATAYAMA N. & SATOH S., "The SR-tree: An Index Structure for High-Dimensional Nearest Neighbor Queries," Proceedins of the ACM SIGMOD international conference on Management of data (Tucson, AZ USA), 11 May 1997 (11.05.97), pp.369-380	1-29
A	BRIN S., "Near Neighbor Search in Large Metric Spaces." Proceedings of the 21st International Conference on Very	1-29

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
 "A" document defining the general state of the art which is not considered to be of particular relevance
 "E" earlier document but published on or after the international filing date

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/09079

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>Large Data Bases, (Zurich, Switzerland), 11 September, 1995 (11.09.95), pp.574-584</p> <p>BERCHTOLD S. et. al., "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality." Proceedings of ACM SIGMOD international conference on Management of data (Seattle, WA USA), 01 June 1998(01.06.98), pp.142-153</p>	1-29

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G06F17/30

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G06F17/30

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2001年
 日本国登録実用新案公報 1994-2001年
 日本国実用新案登録公報 1996-2001年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICSTファイル(JOIS) (ベクトル、ノルム、インデックス、索引、高次元、データベース)
 WPI, INSPEC (DIALOG) (vector, near neighbor, index, search, retrieve, high dimension, database)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	LEE J. Y. et. al. "A fast retrieval method for image features." TENCON 99. Proceedings of the IEEE Region 10 Conference, Vol. 1, 15 Sept. 1999 (15. 09. 99), pp. 577-580	1-29
A	NENE S. A. & NAYAR S. K. "A simple algorithm for nearest neighbor search in high dimensions." IEEE Transactions on Pattern Analysis and Intelligence, Vol. 19, No. 9, Sept. 1997 (09. 97), pp. 989-1003	1-29

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献
 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

26. 01. 00

国際調査報告の発送日

06.02.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

高瀬 勤



51 9069

電話番号 03-3581-1101 内線 3560

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	鷲澤「k-Nearest Neighbor探索の高速計算アルゴリズム」画像電子学会誌, Vol. 27, No. 5, (日), 25. 10月. 1998(25. 10. 98), pp. 513-520	1 - 2 9
A	KATAYAMA N. & SATOH S. "The SR-tree: An Index Structure for High-Dimensional Nearest Neighbor Queries." Proceedins of the ACM SIGMOD international conference on Management of data (Tucson, AZ USA), 11 May 1997(11. 05. 97) pp. 369-380	1 - 2 9
A	BRIN S. "Near Neighbor Search in Large Metric Spaces. Proceedings of the 21st International Conference on Very Large Data Bases, (Zurich, Switzerland), 11 Sept. 1995(11. 09. 95) pp. 574-584	1 - 2 9
A	BERCHTOLD S. et. al. "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality." Proceedings of ACM SIGMOD international conference on Management of data (Seattle, WA USA), 1 June 1998(01. 06. 98), pp. 142-153	1 - 2 9

国際調査報告

(法8条、法施行規則第40、41条)
〔PCT18条、PCT規則43、44〕

出願人又は代理人 の書類記号 P 6 2 - 0 0 0 9	今後の手続きについては、国際調査報告の送付通知様式(PCT/ISA/220) 及び下記5を参照すること。	
国際出願番号 PCT/J P 0 0 / 0 9 0 7 9	国際出願日 (日.月.年) 2 1 . 1 2 . 0 0	優先日 (日.月.年) 2 1 . 1 2 . 9 9
出願人(氏名又は名称) 松下電器産業株式会社		

国際調査機関が作成したこの国際調査報告を法施行規則第41条(PCT18条)の規定に従い出願人に送付する。
この写しは国際事務局にも送付される。

この国際調査報告は、全部で 3 ページである。

☐ この調査報告に引用された先行技術文献の写しも添付されている。

1. 国際調査報告の基礎

a. 言語は、下記に示す場合を除くほか、この国際出願がされたものに基づき国際調査を行った。

☐ この国際調査機関に提出された国際出願の翻訳文に基づき国際調査を行った。

b. この国際出願は、ヌクレオチド又はアミノ酸配列を含んでおり、次の配列表に基づき国際調査を行った。

☐ この国際出願に含まれる書面による配列表

☐ この国際出願と共に提出されたフレキシブルディスクによる配列表

☐ 出願後に、この国際調査機関に提出された書面による配列表

☐ 出願後に、この国際調査機関に提出されたフレキシブルディスクによる配列表

☐ 出願後に提出した書面による配列表が出願時における国際出願の開示の範囲を超える事項を含まない旨の陳述書の提出があった。

☐ 書面による配列表に記載した配列とフレキシブルディスクによる配列表に記載した配列が同一である旨の陳述書の提出があった。

2. ☐ 請求の範囲の一部の調査ができない(第I欄参照)。

3. ☐ 発明の単一性が欠如している(第II欄参照)。

4. 発明の名称は ☒ 出願人が提出したものを承認する。

☐ 次に示すように国際調査機関が作成した。

5. 要約は ☒ 出願人が提出したものを承認する。

☐ 第III欄に示されているように、法施行規則第47条(PCT規則38.2(b))の規定により国際調査機関が作成した。出願人は、この国際調査報告の発送の日から1カ月以内にこの国際調査機関に意見を提出することができる。

6. 要約書とともに公表される図は、

第 1 図とする。 ☒ 出願人が示したとおりである。

☐ なし

☐ 出願人は図を示さなかった。

☐ 本図は発明の特徴を一層よく表している。

THIS PAGE BLANK (USPTO)

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G06F17/30

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G06F17/30

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2001年
 日本国登録実用新案公報 1994-2001年
 日本国実用新案登録公報 1996-2001年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICSTファイル(JOIS) (ベクトル、ノルム、インデックス、索引、高次元、データベース)
 WPI, INSPEC(DIALOG) (vector, near neighbor, index, search, retrieve, high dimension, database)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	LEE J.Y. et.al. "A fast retrieval method for image features." TENCON 99. Proceedings of the IEEE Region 10 Conference, Vol. 1, 15 Sept. 1999 (15. 09. 99), pp. 577-580	1-29
A	NENE S.A. & NAYAR S.K. "A simple algorithm for nearest neighbor search in high dimensions." IEEE Transactions on Pattern Analysis and Intelligence, Vol. 19, No. 9, Sept. 1997 (09. 97), pp. 989-1003	1-29

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献
 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

26. 01. 00

国際調査報告の発送日

06.02.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
 郵便番号 100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

高瀬 勤



5 L

9069

電話番号 03-3581-1101 内線 3560

THIS PAGE BLANK (USPTO)

C (続き) . 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	鷲澤 「k-Nearest Neighbor探索の高速計算アルゴリズム」 画像電子学会誌, Vol. 27, No. 5, (日), 25. 10月. 1998 (25. 10. 98), pp. 513-520	1 - 2 9
A	KATAYAMA N. & SATOH S. "The SR-tree: An Index Structure for High-Dimensional Nearest Neighbor Queries." Proceedings of the ACM SIGMOD international conference on Management of data (Tucson, AZ USA), 11 May 1997 (11. 05. 97) pp. 369-380	1 - 2 9
A	BRIN S. "Near Neighbor Search in Large Metric Spaces." Proceedings of the 21st International Conference on Very Large Data Bases, (Zurich, Switzerland), 11 Sept. 1995 (11. 09. 95) pp. 574-584	1 - 2 9
A	BERCHTOLD S. et. al. "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality." Proceedings of ACM SIGMOD international conference on Management of data (Seattle, WA USA), 1 June 1998 (01. 06. 98), pp. 142-153	1 - 2 9

THIS PAGE BLANK (USPTO)